Equivocations, Exponents and Second-Order Coding Rates under Various Rényi Information Measures

Masahito Hayashi†, *Senior Member, IEEE* and Vincent Y. F. Tan‡, *Senior Member, IEEE*

Abstract

We evaluate the asymptotics of equivocations, their exponents as well as their second-order coding rates under various Rényi information measures. Specifically, we consider the effect of applying a hash function on a source and we quantify the level of non-uniformity and dependence of the compressed source from another correlated source when the number of copies of the sources is large. Unlike previous works that use Shannon information measures to quantify randomness, information or uniformity, we define our security measures in terms of a more general class of information measures—the Rényi information measures and their Gallager-type counterparts. A special case of these Rényi information measure is the class of Shannon information measures. We prove tight asymptotic results for the security measures and their exponential rates of decay. We also prove bounds on the second-order asymptotics and show that these bounds match when the magnitudes of the second-order coding rates are large. We do so by establishing new classes non-asymptotic bounds on the equivocation and evaluating these bounds using various probabilistic limit theorems asymptotically.

Index Terms

Information-theoretic security, Equivocation, Conditional Rényi entropies, Rényi divergence, Sibson's mutual information, Arimoto's mutual information, Error exponents, Secrecy Exponents, Second-order coding rates

I. Introduction

Consider the situation where we are given n independent and identically distributed (i.i.d.) copies of a joint source (A^n, E^n). One of the central tasks in information-theoretic security is to understand the effect of applying a hash function [1] (binning operator) f on A^n. This hash function is used to ensure that the compressed source $f(A^n)$ is almost uniform on its alphabet and also almost independent of another discrete memoryless source E^n. Mathematically, we want to understand the deviation of $f(A^n) \in \{1, \ldots, \lceil \mathrm{e}^{nR} \rceil\}$ from the uniform distribution on the same support $P_{\mathrm{mix}, f(\mathcal{A}^n)}$ and the level of remaining dependence between $f(A^n)$ and a correlated source E^n. These two criteria can be described by equivocation measures. Traditionally in information-theoretic security [2], [3], equivocation is measured in terms of the Shannon-type quantities such as the Shannon entropy, relative entropy (Kullback-Leibler divergence), and mutual information. In particular, it is common to design f such that the following is small for any rate R:

$$D(P_{f(A^n),E^n} \| P_{\mathrm{mix}, f(\mathcal{A}^n)} \times P_{E^n}) = nR - H(f(A^n)|E^n). \tag{1}$$

Clearly if the above quantity is small in some sense, the message $f(A^n)$ is close to uniform and almost independent of E^n, two desirable traits of a hash function for security applications.

A. Motivations

A novel feature of this paper is that we depart from using Shannon information measures to quantify randomness and independence. It is known that the Shannon entropy H or the relative entropy D are special cases of a larger family of information measures known as Rényi information measures, denoted as H_{1+s} and D_{1+s} for $s \in \mathbb{R}$. Thus, as expounded by Iwamoto and Shikata [4], we can quantify equivocation using these measures, gaining

† M. Hayashi is with the Graduate School of Mathematics, Nagoya University, and the Centre for Quantum Technologies (CQT), National University of Singapore (Email: masahito@math.nagoya-u.ac.jp).

‡ V. .Y. F. Tan is with the Department of Electrical and Computer Engineering and the Department of Mathematics, National University of Singapore (Email: vtan@nus.edu.sg).

This paper was presented in part at the 2015 International Symposium on Information Theory in Hong Kong.

deeper insights into the fundamental limits of information leakage under the effect of hash functions. There may also be a possibility of the optimal key generation rate changing when we use alternative information measures. In addition, in the study of cryptography and quantum key distribution (QKD), the Rényi entropy of order 2 [5] (or collision entropy) $H_2(A|P_A) := -\log\sum_{a\in\mathcal{A}} P_A(a)^2$ and the min-entropy $H_{\min}(A|P_A) := -\log\max_{a\in\mathcal{A}} P_A(a)$ play important roles in quantifying randomness. A case in point is the *leftover hash lemma* [6]–[8]. Another motivation stems from the recent study of *overcoming weak expectations* by Dodis and Yu [9] where cryptographic primitives are based on weak secrets, in which the only information about the secret is some fraction of min-entropy. The authors in [9] provided bounds on the weak expectation $\mathbb{E}f(Y)$ of some function f of a random variable Y in terms of the min-entropy and the Rényi entropy of order 2. In a follow-on paper by Yao and Li [10], this study was generalized to Rényi entropies of general orders. Finally, in the study of *secure authentication codes* (or *A-codes* in short), which is one of the most fundamental cryptographic protocols in information-theoretic cryptography, Shikata [11] quantified lengths of secret keys in terms of Rényi entropies of general orders. Motivated by these studies, the authors opine that it is of interest to study the performance of hashing under these generalized families of entropies (generalized uncertainty measures) and divergences (generalized distance measures).

B. Main Contributions

We consider three asymptotic settings—the asymptotics of Rényi-type security measures, its exponential decay and a certain second-order behavior.

1) First, we characterize the asymptotic behavior of the security measure

$$D_{1+s}(P_{f(A^n),E^n}\|P_{\mathrm{mix},f(\mathcal{A}^n)}\times P_{E^n}) \tag{2}$$

for a fixed rate $R=\frac{1}{n}\log\|f\|$ where $\|f\|:=|f(\mathcal{A}^n)|$ is the cardinality of the range of a hash function f. The function f will be taken to be a *random* hash function as we will explain and motivate later. Further, as we shall see in Section II-B, the quantity in (2) is closely related to the equivocation [12]. In Section III (particularly in Corollary 1 therein), we show that if we measure security using D_{1+s} with $s>0$, the fundamental limits of key generation rates change relative to those for traditional Shannon-type measures D_1. The security measure in (2) quantifies the deviation of the hashed or compressed random variable $f(A^n)$ from the uniform distribution and also its remaining dependence from a correlated random variable E^n.

2) We are also interested in the speed of the exponential decay of (2) given a fixed rate R. That is, we are interested in the asymptotic behavior of

$$\frac{1}{n}\log D_{1+s}(P_{f(A^n),E^n}\|P_{\mathrm{mix},f(\mathcal{A}^n)}\times P_{E^n}). \tag{3}$$

This is likened to *error exponent* or *reliability function* analysis in classical information theory [13], [14]. We study this in Section IV.

3) Finally, in Section V, we also study the second-order asymptotics [15], [16] of the decay of D_{1+s} with the blocklength, i.e., the asymptotic behavior of

$$\frac{1}{\sqrt{n}}D_{1+s}(P_{f(A^n),E^n}\|P_{\mathrm{mix},f(\mathcal{A}^n)}\times P_{E^n}),\qquad \text{and} \tag{4}$$

$$\frac{1}{\sqrt{n}}\log D_{1+s}(P_{f(A^n),E^n}\|P_{\mathrm{mix},f(\mathcal{A}^n)}\times P_{E^n}). \tag{5}$$

where the number of compressed symbols (size of the hash function) $\|f\|$ equals $\mathrm{e}^{nR+\sqrt{n}L}$ for some first-order rate R (usually the conditional Rényi entropy) and second-order rate $L\in\mathbb{R}$. For some cases (Rényi parameter less than one) where we cannot exactly determine the tight second-order asymptotics (i.e., the upper and lower bounds do not match), we study the asymptotic behavior of (2) when the second-order rate L tends to $+\infty$ or $-\infty$. In this case, the upper and lower bounds match up to and including a term quadratic in L.

As we mentioned earlier, we will regard f as a random hash function in the sequel. That is, it is randomly selected depending on a random variable $X_n\in\mathcal{X}_n$ that is available to all parties and is also independent of all other random variables. This random variable has distribution P_{X_n}. See Fig. 1. To further elaborate, instead of the

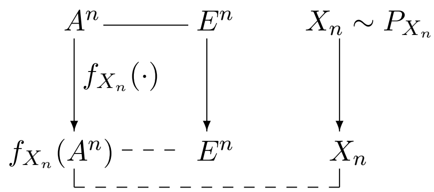

Fig. 1. Illustration of applying a hash function f_{X_n} on the source A^n. Common randomness X_n, independent of a correlated source E^n, is available to all parties and it determines the hash function f_{X_n}. We would like $f_{X_n}(A^n)$ to be uniform on its support $\{1, \ldots, \|f_{X_n}\|\}$ and almost independent of E^n in the sense of ensuring that quantity in (2) is small. We examine (2) under different asymptotic regimes such as the equivocation, the exponential behavior (3), and the second-order asymptotics (4)–(5).

the Rényi divergences in (2)–(5), for the purposes of asserting the existence of a particular function f with some desired properties (cf. the random selection argument), we consider the quantity

$$D_{1+s}^{(n)} := D_{1+s}(P_{f_{X_n}(A^n),E^n,X_n} \| P_{\mathrm{mix}, f_{X_n}(\mathcal{A}^n)} \times P_{E^n} \times P_{X_n}). \tag{6}$$

Here, we note that f_{X_n} is a random hash function (to be defined precisely in Definition 1) and $\|f_{X_n}(\mathcal{A}^n)\|$ is a constant random variable, i.e., it does not depend on the realization of X_n. Even though $D_{1+s}^{(n)}$ in (6) is not an expectation of any quantity of interest, $\exp\big((1+s)D_{1+s}^{(n)}\big)$ is the expectation of

$$\tilde{D}_{1+s}^{(n)}(x_n) := \exp\big((1+s)D_{1+s}(P_{f_{x_n}(A^n),E^n} \| P_{\mathrm{mix}, f_{x_n}(\mathcal{A}^n)} \times P_{E^n})\big), \tag{7}$$

where the probability of observing x_n is $P_{X_n}(x_n)$. Thus by a random selection argument, if the former is less than $\varepsilon > 0$, there exist an $x_n^* \in \mathcal{X}_n$, indexing a deterministic protocol $f_{x_n^*}$, such that $\tilde{D}_{1+s}^{(n)}(x_n^*)$ is also less than ε. When $s = 0$, the expectation of quantities in (2)–(5) under the common randomness X_n generating a universal$_2$ hash function $f_{X_n}(\cdot)$ is equivalent to the quantity in (6) but for $s \neq 0$, they are, in general, different. In the sequel, we adopt the latter criterion in (6) to simplify the presentation of the results.

We believe the results contained herein may serve as logical starting points to derive *tight* exponential error bounds and second-order coding rates for the wiretap channel [12] (as was done in [17], [18]) and other information-theoretic security problems such as the secret key agreement [19] (as was done in [20], [21]) problem. The leakage rates for these problems may be measured using traditional Shannon information measures or Rényi information measures (or their Gallager-type counterparts). Here, we are only concerned with the secrecy requirement rather than both the secrecy and reliability requirements of the wiretap problem. The reliability requirement can be handled using, by now, standard error exponent analyses [13], [14].

C. *Related Works*

In [17], [22], Hayashi generalized and strengthened the seminal privacy amplification analyses of Bennett *et al.* [7], Renner [23] and Renner and Wolf [24] to obtain exponential error bounds for the leakage rate of the discrete memoryless wiretap channel and the secrey key agreement problems [2], [3], [12], [19]. The leakage rate was measured by the mutual information $I(A \wedge E|P_{AE})$ and the variational (or trace) distance $\|P_{AE} - P_A \times P_E\|_1$. The exact exponents for the variational distance are, by now, well known [18], [25].

However, the results concerning the exponential decay of the leakage rate quantified via the mutual information contained in [17], and further generalized to other setings in [20], [21], are only *achievability* results (i.e., lower bounds on the exponents). The *converse* has been open for some time. The present contribution, though not focusing on the wiretap channel or any specific information-theoretic security problem, derives *tight* exponential bounds for a generalization of the Shannon information measures, namely the family of Rényi information measures. In the process, we obtain a tight result for the exponential leakage rate for the mutual information, thus resolving the converse part that was open in [17], [20], [21]. As a by-product, for some range of the Rényi parameter, we also obtain tight exponents for security measures defined using the Rényi divergence under various hash functions.

Hayashi and Tsurumaru [26] proposed an efficient construction of hash functions for the purpose of privacy amplification with less random seeds, thus potentially realizing the system in Fig. 1 with less random resources. Other works along the lines of deriving exponential error bounds for information-theoretic security problems include those by Hou and Kramer [27], [28], Pierrot and Bloch [29], Bloch and Laneman [30], Han *et al.* [31] and Parizi

and Telatar [32]. After the present work was submitted, Parizi, Telatar and Merhav [33] proved ensemble tight exponential error bounds for the wiretap channel by appealing to type counting methods and channel resolvability arguments.

D. Paper Organization

The rest of the paper is organized as follows: In Section II, we state the relevant preliminaries and define relevant information measures and security criteria for understanding the rest of the paper. In Section III, we state our results for the asymptotics of the equivocation. In Section IV, we state our results for the exponential behavior of the Rényi-type security criteria. In Section V, we state our results for the second-order asymptotics of the equivocation. We also consider the case where the magnitudes of the second-order rates are large. These are proved using novel one-shot bounds which are stated in Section VI. The proofs of the asymptotic results are provided in Section VII. We conclude the paper in Section VIII by summarizing our key contributions and stating avenues for further investigations. The proofs of the one-shot bounds are rather technical and are thus relegated to the Appendices.

II. Preliminaries and Information Measures

A. Basic Shannon and Rényi Information Quantities

We now introduce some information measures that generalize Shannon's information measures. Fix a normalized distribution $P_A \in \mathcal{P}(\mathcal{A})$ and a non-negative measure (a non-negative vector but not necessarily summing to one) $Q_A \in \bar{\mathcal{P}}(\mathcal{A})$ supported on a finite set $\mathcal{A}$. Then the *relative entropy* and the *Rényi divergence of order* $1+s$ are respectively defined as

$$D(P_A \| Q_A) := \sum_{a \in \mathcal{A}} P_A(a) \log \frac{P_A(a)}{Q_A(a)} \tag{8}$$

$$D_{1+s}(P_A \| Q_A) := \frac{1}{s} \log \sum_{a \in \mathcal{A}} P_A(a)^{1+s} Q_A(a)^{-s}, \tag{9}$$

where throughout, $\log$ is to the natural base e. It is known that $\lim_{s \to 0} D_{1+s}(P_A \| Q_A) = D(P_A \| Q_A)$ so a special case of the Rényi divergence is the usual relative entropy. It is known that the map $s \mapsto sD_{1+s}(P_A \| Q_A)$ is concave in $s \in \mathbb{R}$ and hence $D_{1+s}(P_A \| Q_A)$ is monotonically increasing for $s \in \mathbb{R}$. Furthermore, the following *data processing or information processing inequalities* for Rényi divergences hold for $s \in [-1, 1]$,

$$D(P_A W \| Q_A W) \le D(P_A \| Q_A) \tag{10}$$

$$D_{1+s}(P_A W \| Q_A W) \le D_{1+s}(P_A \| Q_A). \tag{11}$$

Here $W : \mathcal{A} \to \mathcal{B}$ is any stochastic matrix (channel) and $P_A W(b) := \sum_a W(b|a) P_A(a)$ is the output distribution induced by W and P_A.

We use $P_{\mathrm{mix},\mathcal{A}}$ to denote the uniform distribution on $\mathcal{A}$. We also introduce conditional entropies on the joint alphabet $\mathcal{A} \times \mathcal{E}$. If P_{AE} is a distribution on $\mathcal{A} \times \mathcal{E}$, the *conditional entropy* and the *conditional Rényi entropy of order* $1+s$ *relative to another normalized distribution* Q_E *on* $\mathcal{E}$ as

$$H(A|E|P_{AE} \| Q_E) := -D(P_{AE} \| I_A \times Q_E), \tag{12}$$

$$H_{1+s}(A|E|P_{AE} \| Q_E) := -D_{1+s}(P_{AE} \| I_A \times Q_E). \tag{13}$$

Here $I_A(a) = 1$ for each $a \in \mathcal{A}$ and it is known that $\lim_{s \to 0} H_{1+s}(A|E|P_{AE} \| Q_E) = H(A|E|P_{AE} \| Q_E)$. If $Q_E = P_E$, we simplify the notation and denote the *conditional entropy* and the *conditional Rényi entropy of order* $1+s$ as

$$H(A|E|P_{AE}) := H(A|E|P_{AE} \| P_E) = -\sum_e P_E(e) \sum_a P_{A|E}(a|e) \log P_{A|E}(a|e) \tag{14}$$

$$H_{1+s}(A|E|P_{AE}) := H_{1+s}(A|E|P_{AE} \| P_E) = -\frac{1}{s} \log \sum_e P_E(e) \sum_a P_{A|E}(a|e)^{1+s}. \tag{15}$$

The function $s \mapsto sH_{1+s}(A|E|P_{AE})$ is concave, and $H_{1+s}(A|E|P_{AE}\|Q_E)$ is monotonically decreasing on $(0,\infty)$ and $(-\infty,0)$.

We are also interested in the so-called *Gallager form* of the conditional Rényi entropy for a joint distribution $P_{AE} \in \mathcal{P}(\mathcal{A}\times\mathcal{E})$:

$$H^{\uparrow}_{1+s}(A|E|P_{AE}) := -\frac{1+s}{s}\log\sum_e\left(\sum_a P_{AE}(a,e)^{1+s}\right)^{\frac{1}{1+s}}. \tag{16}$$

By defining the familiar *Gallager function* [13], [34] (parametrized slightly differently)

$$\phi(s|A|E|P_{AE}) := \log\sum_e\left(\sum_a P_{AE}(a,e)^{\frac{1}{1-s}}\right)^{1-s} \tag{17}$$

we can express (16) as

$$H^{\uparrow}_{1+s}(A|E|P_{AE}) = -\frac{1+s}{s}\phi\left(\frac{s}{1+s}\middle|A|E|P_{AE}\right), \tag{18}$$

thus (loosely) justifying the nomenclature "Gallager form" of the conditional Rényi entropy in (16). The quantities H_{1+s} and $H^{\uparrow}_{1+s}$ can be shown to be related as follows:

$$\max_{Q_E\in\mathcal{P}(\mathcal{E})} H_{1+s}(A|E|P_{AE}\|Q_E) = H^{\uparrow}_{1+s}(A|E|P_{AE}) \tag{19}$$

for $s \in [-1,\infty)\setminus\{0\}$. The maximum on the left-hand-side is attained for the tilted distribution

$$Q_E(e) = \frac{(\sum_a P_{AE}(a,e)^{1+s})^{\frac{1}{1+s}}}{\sum_e(\sum_a P_{AE}(a,e)^{1+s})^{\frac{1}{1+s}}}. \tag{20}$$

The map $s \to sH^{\uparrow}_{1+s}(A|E|P_{AE})$ is concave and the map $s \mapsto H^{\uparrow}_{1+s}(A|E|P_{AE})$ is monotonically decreasing for $s \in (-1,\infty)$. It can be shown by L'Hôpital's rule that

$$\lim_{s\to 0} H^{\uparrow}_{1+s}(A|E|P_{AE}) = H(A|E|P_{AE}). \tag{21}$$

Thus, we regard $H^{\uparrow}_1(A|E|P_{AE})$ as $H(A|E|P_{AE})$, i.e., for Rényi parameter $\alpha = 1+s = 1$, the conditional Rényi entropy and its Gallager form coincide. We also find it useful to consider a *two-parameter family* of the conditional Rényi entropy:

$$H_{1+s|1+t}(A|E|P_{AE}) := -\frac{1+t}{s}\log\sum_e P_E(e)\left(\sum_a P_{A|E}(a|e)^{1+s}\right)^{\frac{1}{1+t}}. \tag{22}$$

Clearly,

$$H_{1+s|1+s}(A|E|P_{AE}) = H^{\uparrow}_{1+s}(A|E|P_{AE}) \tag{23}$$

so two-parameter conditional Rényi entropy is a generalization of the Gallager form of the conditional Rényi entropy in (16).

For a fixed joint source P_{AE} define

$$\hat{R}_s := \frac{\mathrm{d}}{\mathrm{d}t}\, tH_{1+t}(A|E|P_{AE})\Big|_{t=s}, \quad \text{and} \tag{24}$$

$$\hat{R}^{\uparrow}_s := \frac{\mathrm{d}}{\mathrm{d}t}\, tH^{\uparrow}_{1+t}(A|E|P_{AE})\Big|_{t=s}. \tag{25}$$

We note that $\hat{R}_s$ and $\hat{R}^{\uparrow}_s$ are monotonically non-increasing in s because the functions $t \mapsto tH_{1+t}(A|E|P_{AE})$ and $t \mapsto tH^{\uparrow}_{1+t}(A|E|P_{AE})$ are concave. The fact that $t \mapsto -tH^{\uparrow}_{1+t}(A|E|P_{AE})$ is convex is because the maximum of convex functions is convex; cf. (19). Furthermore, both $\hat{R}_s$ and $\hat{R}^{\uparrow}_s$ are non-negative by direct evaluation of the derivatives and noting that $\log P_{A|E}(a|e) \le 0$. We assume, henceforth, that the source P_{AE} satisfies the conditions that $t \mapsto tH_{1+t}(A|E|P_{AE})$ and $t \mapsto tH^{\uparrow}_{1+t}(A|E|P_{AE})$ are both *strictly concave* so $\hat{R}_s$ and $\hat{R}^{\uparrow}_s$ are both *monotonically decreasing* in s.

The Rényi entropies can be shown to satisfy a form of *data processing inequality*. In particular if $f : \mathcal{A} \to \mathcal{M}$ is any function on the set $\mathcal{A}$, we have

$$H(f(A)|E|P_{AE}) \le H(A|E|P_{AE}), \tag{26}$$
$$H_{1+s}(f(A)|E|P_{AE}) \le H_{1+s}(A|E|P_{AE}), \tag{27}$$
$$H^{\uparrow}_{1+s}(f(A)|E|P_{AE}) \le H^{\uparrow}_{1+s}(A|E|P_{AE}). \tag{28}$$

Inequalities (27) and (28) hold true for all $s > -1$. These inequalities say that processing the random variable A cannot increase its randomness measured under any of the above conditional Rényi entropies.

B. Rényi Security Criteria

Now, we introduce various criteria that measure independence and uniformity *jointly*. The *mutual information* is

$$I(A \wedge E|P_{AE}) := D(P_{AE}\|P_A \times P_E). \tag{29}$$

This, together with its normalized version, has been traditionally used as measure of dependence in classical information-theoretic security [2], [3], going back to the seminal work of Wyner [12] for the wiretap channel. It was also used by Ahlswede and Csiszár for the secret key agreement problem [19]. However, it does not guarantee approximate uniformity of the source P_A on $\mathcal{A}$. Thus, we introduce the *modified mutual information*

$$C(A|E|P_{AE}) := D(P_{AE}\|P_{\mathrm{mix},\mathcal{A}} \times P_E) \tag{30}$$
$$= \log|\mathcal{A}| - H(A|E|P_{AE}). \tag{31}$$

This quantity was also considered by Csiszár and Narayan [35, Eq. (6)] in their work on secrecy capacities. An axiomatic justification of $C(A|E|P_{AE})$ was provided recently by Hayashi [36, Thm. 8]. The modified mutual information $C(A|E|P_{AE})$ clearly satisfies

$$C(A|E|P_{AE}) = I(A \wedge E|P_{AE}) + D(P_A\|P_{\mathrm{mix},\mathcal{A}}). \tag{32}$$

Hence, if $C(A|E|P_{AE})$ is small, A is approximately independent of E *and* A is approximately uniform on its alphabet, desirable properties in information-theoretic security. We may further generalize the modified mutual information by considering *Rényi information measures*, introduced in Section II-A, as follows:

$$C_{1+s}(A|E|P_{AE}) := D_{1+s}(P_{AE}\|P_{\mathrm{mix},\mathcal{A}} \times P_E) \tag{33}$$
$$= \log|\mathcal{A}| - H_{1+s}(A|E|P_{AE}). \tag{34}$$

This can be relaxed to give yet another security measure—the *Gallager-form of the modified mutual information*:

$$C^{\uparrow}_{1+s}(A|E|P_{AE}) := \min_{Q_E \in \mathcal{P}(\mathcal{E})} D_{1+s}(P_{AE}\|P_{\mathrm{mix},\mathcal{A}} \times Q_E) \tag{35}$$
$$= \log|\mathcal{A}| - H^{\uparrow}_{1+s}(A|E|P_{AE}). \tag{36}$$

We characterize these quantities asymptotically when $(A, E) \equiv (f(A^n), E^n)$ for some (classes of) hash functions $f(\cdot)$. The quantities H_{1+s} and $H^{\uparrow}_{1+s}$ can be regarded as equivocations [12] so C_{1+s} and $C^{\uparrow}_{1+s}$ are the negative of the equivocations up to a shift. We work with C_{1+s} and $C^{\uparrow}_{1+s}$ in the rest of the paper as they are more convenient and they admit the interpretation as *security criteria*.

C. Decomposition of the Rényi Security Criteria into Mutual Information and Divergence Terms

We note that for any $s \ge -1$, $C_{1+s}(A|E|P_{AE}) = 0$ if and only if $P_{AE} = P_{\mathrm{mix},\mathcal{A}} \times P_E$ or equivalently, A is uniform on $\mathcal{A}$ and statistically independent of E. This is because $D_{1+s}(P\|Q)$ is a divergence so $D_{1+s}(P\|Q) = 0$ if and only if $P = Q$ [37]. The same is true for the case $C^{\uparrow}_{1+s}(A|E|P_{AE}) = 0$. From this observation, we see that $C_{1+s}(A|E|P_{AE})$ and $C^{\uparrow}_{1+s}(A|E|P_{AE})$ also measure how close the source or "key" A is to uniform and how secure A is from an adversary E. Thus the quantities we consider are generalizations of the standard security measure $C(A|E|P_{AE})$ in (32) and measure uniformity and security in a different way.

More quantitatively, one may wonder whether the security criteria $C_{1+s}(A|E|P_{AE})$ and $C^{\uparrow}_{1+s}(A|E|P_{AE})$ admit a decomposition into "mutual information" and "divergence" terms and similar to (32). We first consider $C_{1+s}(A|E|P_{AE})$. Define $g_s(a) := \sum_e P_{AE}(a,e)^{1+s} P_E(e)^{-s}$. We then see from the definition of the Rényi divergence of order $(1+s)$ that

$$\begin{aligned}
&\mathrm{e}^{sD_{1+s}(P_{AE}\|P_{\mathrm{mix},\mathcal{A}}\times P_E)}\\
&= \left(\frac{1}{|\mathcal{A}|}\right)^{-s} \sum_a \left(\sum_e P_{AE}(a,e)^{1+s} P_E(e)^{-s} \right) && (37)\\
&= \left(\frac{1}{|\mathcal{A}|}\right)^{-s} \sum_a \left(g_s(a)^{\frac{1}{1+s}} \right)^{1+s} && (38)\\
&= \left(\frac{1}{|\mathcal{A}|}\right)^{-s} \left(\sum_{a'} g_s(a')^{\frac{1}{1+s}} \right)^{1+s} \sum_a \left(\frac{g_s(a)^{\frac{1}{1+s}}}{\sum_{a'} g_s(a')^{\frac{1}{1+s}}} \right)^{1+s}. && (39)
\end{aligned}$$

As a result, one has

$$D_{1+s}(P_{AE}\|P_{\mathrm{mix},\mathcal{A}}\times P_E) = \frac{1+s}{s} \log \sum_{a'} g_s(a')^{\frac{1}{1+s}} + \frac{1}{s} \log \sum_a \left(\frac{g_s(a)^{\frac{1}{1+s}}}{\sum_{a'} g_s(a')^{\frac{1}{1+s}}} \right)^{1+s} + \log|\mathcal{A}|. \tag{40}$$

Invoking the definition of $g_s(a)$, we see that the first term can be rewritten as

$$\begin{aligned}
\frac{1+s}{s} \log \sum_{a'} g_s(a')^{\frac{1}{1+s}} &= \frac{1+s}{s} \log \sum_{a'} P_A(a') \left(\sum_e P_{E|A}(e|a')^{1+s} P_E(e)^{-s} \right)^{\frac{1}{1+s}} && (41)\\
&=: I^{(\mathrm{Sibson})}_{1+s}(E\wedge A|P_{AE}). && (42)
\end{aligned}$$

This is exactly *Sibson's* definition of the order-$(1+s)$ Rényi mutual information [38]. See Verdú's work in [39, Sec. III] for the properties of $I^{(\mathrm{Sibson})}_{1+s}(E\wedge A|P_{AE})$ and a generalization to arbitrary alphabets. See Hayashi's work [40, Sec. II.C] for a generalization of $I^{(\mathrm{Sibson})}_{1+s}(E\wedge A|P_{AE})$ to quantum systems. The work of Tomamichel and Hayashi in [41, Sec. IV.B] provides an operational interpretation of this quantity in the context of composite hypothesis testing. The sum of the second and third terms in (40) form a Rényi divergence of order $(1+s)$. In particular, the second term is the negative Rényi entropy of order $(1+s)$ of the probability mass function $Q_A^{(s)}(a) := g_s(a)^{\frac{1}{1+s}} / \sum_{a'} g_s(a')^{\frac{1}{1+s}}$. Hence,

$$D_{1+s}(P_{AE}\|P_{\mathrm{mix},\mathcal{A}}\times P_E) = I^{(\mathrm{Sibson})}_{1+s}(E\wedge A|P_{AE}) + D_{1+s}\big(Q_A^{(s)}\|P_{\mathrm{mix},\mathcal{A}}\big). \tag{43}$$

Because $Q_A^{(0)} = P_A$, and $\lim_{s\to 0} I^{(\mathrm{Sibson})}_{1+s}(E\wedge A|P_{AE}) = I(A\wedge E|P_{AE})$, the decomposition in (43) is a generalization of (32). Equation (43) is also reminiscent of an information geometric Pythagorean theorem [42] (but for Rényi divergence here). The distribution $Q_A^{(s)}\times P_E$ can be regarded as the D_{1+s}-information projection of P_{AE} onto the set $\{Q_A\times P_E : Q_A\in\mathcal{P}(\mathcal{A})\}$. This was observed in the quantum information context by Sharma and Warsi [43, Lemma 3 in Suppl. Mat.]. They called the relation the *quantum Sibson identity*.

Next we consider the Gallager-form of the modified mutual information $C^{\uparrow}_{1+s}(A|E|P_{AE})$. From (36), it can be seen by adding and subtracting $H_{1+s}(A|P_A)$ that

$$C^{\uparrow}_{1+s}(A|E|P_{AE}) = H_{1+s}(A|P_A) - H^{\uparrow}_{1+s}(A|E|P_{AE}) + D_{1+s}(P_A\|P_{\mathrm{mix},\mathcal{A}}). \tag{44}$$

We recognize that the sum of the first two terms constitutes *Arimoto's* [44] definition of the order-$(1+s)$ Rényi mutual information

$$I^{(\mathrm{Arimoto})}_{1+s}(A\wedge E|P_{AE}) = H_{1+s}(A|P_A) - H^{\uparrow}_{1+s}(A|E|P_{AE}). \tag{45}$$

Since $\lim_{s\to 0} I^{(\mathrm{Arimoto})}_{1+s}(A\wedge E|P_{AE}) = I(A\wedge E|P_{AE})$, the security criterion $C^{\uparrow}_{1+s}(A|E|P_{AE})$ also admits a decomposition similar to (32). See [39, Sec. II.A] for detailed discussions of the properties of $I^{(\mathrm{Arimoto})}_{1+s}(A\wedge E|P_{AE})$.

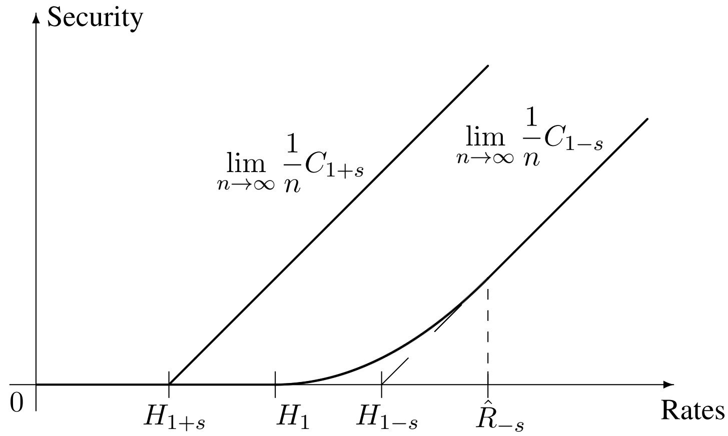

Fig. 2. Schematic showing the relation between the various entropies and the transition rate $\hat{R}_{-s}$ (defined in (24) and (51)). The figure with the Gallager forms of the conditional Rényi entropy $H^{\uparrow}_{1\pm s}$ and $\hat{R}^{\uparrow}_{-s}$ (defined in (25) and (52)) is completely analogous. See Fig. 4.

III. Asymptotics of the Equivocation

In this section we present our results concerning the asymptotic behavior of the equivocation. First we define precisely the notion of hash function. This is a generalization of the definition by Carter and Wegman [1].

Definition 1. *A* random[1] hash function f_X *is a stochastic map from* $\mathcal{A}$ *to* $\mathcal{M} := \{1,\ldots,M\}$*, where* X *denotes a random variable describing its stochastic behavior. An ensemble of random hash functions* f_X *is called an* ϵ-almost universal$_2$ hash function *if it satisfies the following condition: For any* distinct $a_1, a_2 \in \mathcal{A}$*,*

$$\Pr\left(f_X(a_1) = f_X(a_2)\right) \le \frac{\epsilon}{M}. \tag{46}$$

When $\epsilon = 1$*, we simply say that the ensemble of functions is a* universal$_2$ hash function*.*

As an example, if we randomly and uniformly assign each element of $a \in \mathcal{A}$ into one of M bins indexed by $m \in \mathcal{M}$ (i.e., the familiar random binning process introduced by Cover [45]), then $\Pr(f_X(a_1) = f_X(a_1)) = \frac{1}{M}$ so this is a universal$_2$ hash function, and furthermore, (46) is achieved with equality.

Let $|t|^+ = \max\{0, t\}$. The following is our first main result.

Theorem 1 (Asymptotics of the Equivocation)**.** *Let*[2] $M_n = \mathrm{e}^{nR}$*. Assume that* $f_{X_n} : \mathcal{A}^n \to \mathcal{M}_n = \{1,\ldots,M_n\}$ *is a random hash function.*[3] *For any* $s \in [0,1]$*, we have*

$$\lim_{n\to\infty}\frac{1}{n}\inf_{f_{X_n}} C_{1+s}(f_{X_n}(A^n)|E^nX_n|P^n_{AE}\times P_{X_n}) = |R - H_{1+s}(A|E|P_{AE})|^+, \tag{47}$$

$$\lim_{n\to\infty}\frac{1}{n}\inf_{f_{X_n}} C^{\uparrow}_{1+s}(f_{X_n}(A^n)|E^nX_n|P^n_{AE}\times P_{X_n}) = |R - H^{\uparrow}_{1+s}(A|E|P_{AE})|^+. \tag{48}$$

Furthermore, for any $s \in (0,1]$*, we also have*

$$\lim_{n\to\infty}\frac{1}{n}\inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^nX_n|P^n_{AE}\times P_{X_n}) = \begin{cases} R - H_{1-s}(A|E|P_{AE}) & R \ge \hat{R}_{-s} \\ \max_{t\in[0,s]}\frac{t}{s}(R - H_{1-t}(A|E|P_{AE})) & R \le \hat{R}_{-s} \end{cases}, \tag{49}$$

$$\lim_{n\to\infty}\frac{1}{n}\inf_{f_{X_n}} C^{\uparrow}_{1-s}(f_{X_n}(A^n)|E^nX_n|P^n_{AE}\times P_{X_n}) = \begin{cases} R - H^{\uparrow}_{1-s}(A|E|P_{AE}) & R \ge \hat{R}^{\uparrow}_{-s} \\ \max_{t\in[0,s]}\frac{t}{s}(R - H_{1-t|1-s}(A|E|P_{AE})) & R \le \hat{R}^{\uparrow}_{-s} \end{cases}, \tag{50}$$

[1]For brevity, we will sometimes omit the qualifier "random". It is understood, henceforth, that all so-mentioned hash functions are random hash functions.

[2]As is usual in information theory, we ignore the integer effects on the size of the hash function $M_n = \|f\|$ since this is inconsequential asymptotically. This imprecision is also employed in the sequel for notational convenience.

[3]In particular, all the infima in (47)–(50) (as well as similar statements in the sequel) are taken over all f_{X_n} that are random hash functions.

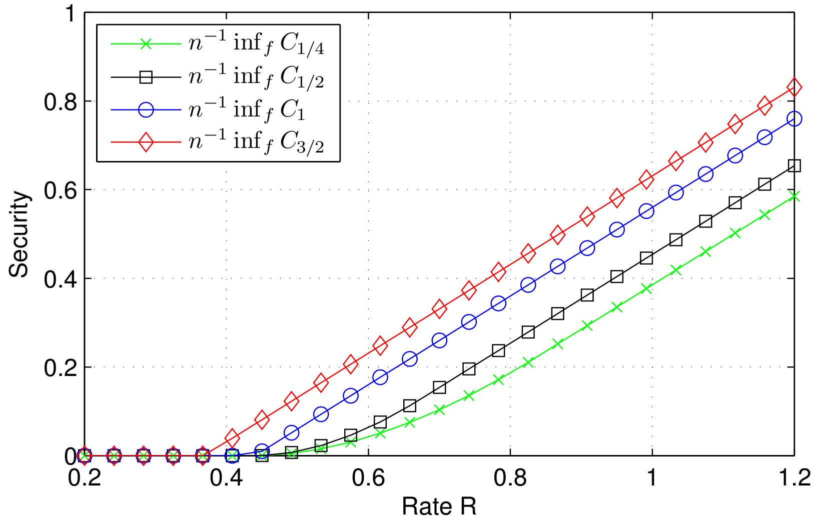

Fig. 3. Illustration of the security measures C_{1+s} and C_{1-s} (for $s \in [0,1]$) in (47) and (49) for the discrete memoryless multiple source P_{AE} where $P_{AE}(0,0) = 0.7$ and $P_{AE}(0,1) = P_{AE}(1,0) = P_{AE}(1,1) = 0.1$.

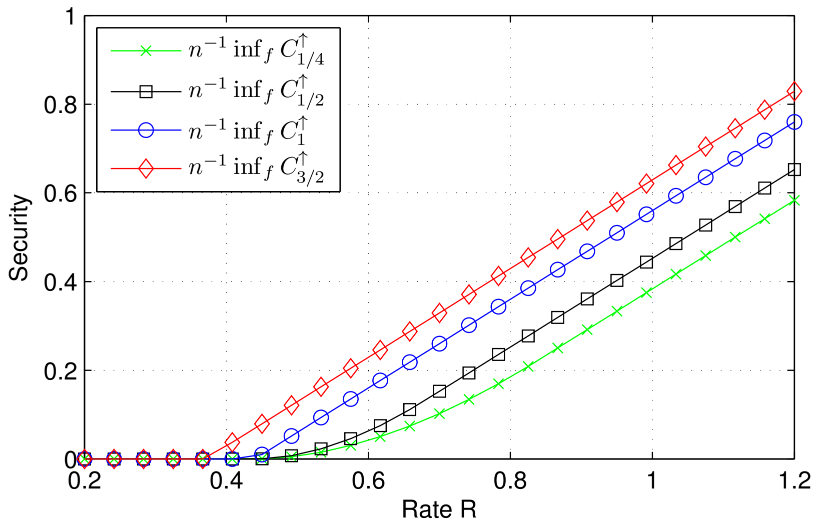

Fig. 4. Illustration of the security measures $C^{\uparrow}_{1+s}$ and $C^{\uparrow}_{1-s}$ (for $s \in [0,1]$) in (48) and (50) respectively for the same source.

where recall that $\hat{R}_{-s}$ *and* $\hat{R}^{\uparrow}_{-s}$ *are defined in* (24) *and* (25) *respectively. (Also see* (51) *and* (52) *for alternative representations.) Furthermore, the infima in* (47)–(50) *are achieved by any sequence of* ϵ*-almost universal*$_2$ *hash functions* f_{X_n} *(where* ϵ *is a fixed positive number).*

This result is proved in Section VII-A. The ideas to prove the direct parts (upper bounds on the leakage rates) for the $s = 0$ cases are contained in previous works such as [17], [20], [21]. All other parts are novel.

We remark that the converse parts (lower bounds) to (47)–(48) hold for all $s \geq 0$ (and not only being upper bounded by 1) owing to the data processing inequalities in (27)–(28). Furthermore, instead of the formulae in (24) and (25), the rates in which the behavior of the security measures change $\hat{R}_{-s}$ and $\hat{R}^{\uparrow}_{-s}$ can also be expressed as

$$\hat{R}_{-s} = \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t}(A|E|P_{AE})\Big|_{t=s}, \quad \text{and} \tag{51}$$

$$\hat{R}^{\uparrow}_{-s} = \frac{\mathrm{d}}{\mathrm{d}t} tH^{\uparrow}_{1-t}(A|E|P_{AE})\Big|_{t=s}. \tag{52}$$

These alternative expressions for $\hat{R}_{-s}$ and $\hat{R}^{\uparrow}_{-s}$ will be useful in the proof of Theorem 1.

The results in (47)–(50) imply that an optimum sequence of hash functions $\{f_{X_n}\}_{n\in\mathbb{N}}$ is such that asymptotically the normalized security measure C_{1+s} and its Gallager-type counterpart $C^{\uparrow}_{1+s}$ increase *linearly* with the rate R if

the rate is larger than the conditional Rényi entropy and its Gallager-type counterpart. However, note that this only holds for the case where R is greater than the analogue of the critical rates, defined in (24)–(25) in the case where the Rényi parameter $\alpha = 1-s$ is less than one. Observe that there is difference in behavior when we consider the other direction, i.e., the quantities C_{1-s} and $C^{\uparrow}_{1-s}$ for $s \in [0,1]$. Below the critical rate, the equivocation no longer increases linearly with R but is nonetheless still convex in R. See Fig. 2 for a schematic of the various rates and the behavior of the equivocations. We numerically calculate the asymptotics of the equivocations in Theorem 1 and display the results in Figs. 3 and 4. The behaviors of the normalized security measure C_{1+s}, C_{1-s} and their Gallager-type counterparts $C^{\uparrow}_{1+s}, C^{\uparrow}_{1-s}$ are similar.

Finally, we examine the optimal (maximum) key generation rates, i.e., the largest rates R for which there exists a sequence of functions from $\mathcal{A}^n$ to $\{1,\ldots,\mathrm{e}^{nR}\}$ such that $\frac{1}{n}C_{1+s}$ or $\frac{1}{n}C^{\uparrow}_{1+s}$ tend to zero as the blocklength grows. We observe from the following corollary that this cutoff rate depends strongly on the sign of s. In particular for $s \in (0,1]$, the cutoff rates are $H_{1+s}(A|E|P_{AE})$ and $H^{\uparrow}_{1+s}(A|E|P_{AE})$ respectively, while for $s \in [-1,0]$, the cutoff rates are both equal to the Shannon conditional entropy $H(A|E|P_{AE})$ independent of s. This difference between the behaviors of the optimal key generation rates depending on the sign of s (also illustrated in Figs. 3 and 4) is somewhat surprising (at least to the authors).

Corollary 1 (Optimal key generation rates)**.** *We have*

$$\sup\left\{R \in \mathbb{R}_+ : \lim_{n\to\infty} \inf_{f:\mathcal{A}^n\to\{1,\ldots,\mathrm{e}^{nR}\}} \frac{C_{1+s}(f(A^n)|E^n|P^n_{AE})}{n} = 0\right\} = \begin{cases} H_{1+s}(A|E|P_{AE}) & \text{if } s \in (0,1] \\ H(A|E|P_{AE}) & \text{if } s \in [-1,0] \end{cases}, \quad (53)$$

$$\sup\left\{R \in \mathbb{R}_+ : \lim_{n\to\infty} \inf_{f:\mathcal{A}^n\to\{1,\ldots,\mathrm{e}^{nR}\}} \frac{C^{\uparrow}_{1+s}(f(A^n)|E^n|P^n_{AE})}{n} = 0\right\} = \begin{cases} H^{\uparrow}_{1+s}(A|E|P_{AE}) & \text{if } s \in (0,1] \\ H(A|E|P_{AE}) & \text{if } s \in [-1,0] \end{cases}. \quad (54)$$

Proof: We only prove the statement for C_{1+s} in (53) since that for $C^{\uparrow}_{1+s}$ in (54) is completely analogous. The case for $s \in (0,1]$ is obvious from (47) in Theorem 1 since the limit is $|R - H_{1+s}(A|E|P_{AE})|^+$. Now, for the case $s \in [-1,0]$, if $R \le H(A|E|P_{AE})$, we know from the monotonically decreasing nature of $H_{1+s}(A|E|P_{AE})$ (in s) that $R - H_{1-t}(A|E|P_{AE})$ is non-positive for $t \in [0,s]$. Thus, referring to (49) in Theorem 1, the optimal t in the optimization $\max_{t\in[0,s]} \frac{t}{s}(R - H_{1-t}(A|E|P_{AE}))$ is attained at $t=0$ and consequently, the optimal objective value is 0. On the other hand, for any $R > H(A|E|P_{AE})$, the optimal $t \in (0,s]$ and so the optimal objective value is (strictly) positive. Thus, for $s \in [-1,0]$, the optimal key generation rate is the Shannon conditional entropy $H(A|E|P_{AE})$. This concludes the proof for (53). ∎

In Section I-A, we alluded to the importance of the collision entropy H_2 in cryptography and QKD. The implication of (53) in Corollary 1 is that if we operate at a hashing rate $R > H_2$ and we employ the security criterion C_2, then there will inevitably be some residual leakage of the source A^n given a hashed version $f(A^n)$ and side-information E^n.

Because of the normalizations of C_{1+s} and $C^{\uparrow}_{1+s}$ by n in (53) and (54), Corollary 1 is analogous to results in the vast majority of the literature in information-theoretic security [2], [3] where the *weak secrecy* criterion is employed. We address the analogue of the *strong secrecy* criterion [46] in Theorem 2 to follow where we not only demand that the unnormalized quantities C_{1+s} and $C^{\uparrow}_{1+s}$ vanish with n, we also demand that they do so exponentially fast and we identify the exponents.

IV. Exponential Behavior of the Security Measures

In this section, we evaluate the exponential rates of decay of the security measures $C_{1\pm s}$ and $C^{\uparrow}_{1\pm s}$ for fixed rates R above an analogue of the critical rate.

Theorem 2 (Exponents of the Equivocation)**.** *Let $M_n = \mathrm{e}^{nR}$. Assume that $f_{X_n} : \mathcal{A}^n \to \mathcal{M}_n = \{1,\ldots,M_n\}$ is a random hash function. For $R \ge \hat{R}_1$ ($\hat{R}_s$ being defined in* (24)*), and any $s \in [0,1]$, we have*

$$\lim_{n\to\infty} -\frac{1}{n}\log \inf_{f_{X_n}} C_{1+s}(f_{X_n}(A^n)|E^n X_n|P^n_{AE}\times P_{X_n}) = \left|\sup_{t\in[s,1)} tH_{1+t}(A|E|P_{AE}) - tR\right|^+, \quad (55)$$

$$\lim_{n\to\infty} -\frac{1}{n}\log \inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^n X_n|P^n_{AE}\times P_{X_n}) = \max_{t\in[0,1]} tH_{1+t}(A|E|P_{AE}) - tR. \quad (56)$$

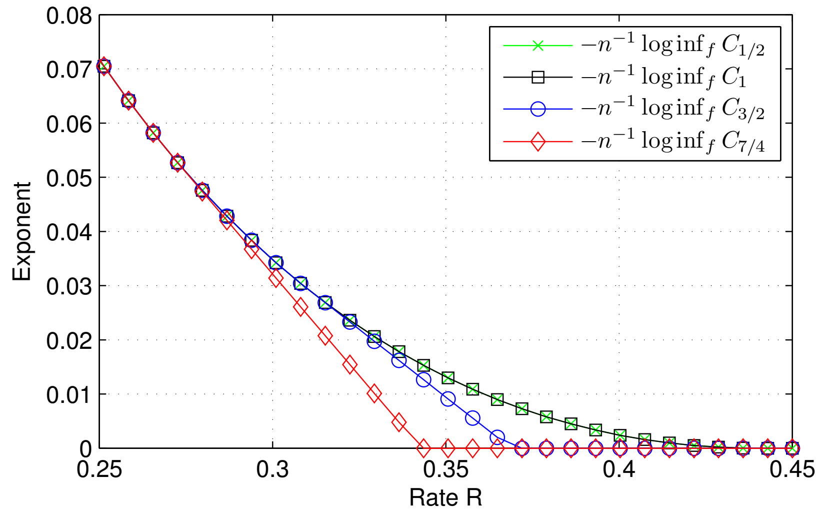

Fig. 5. Illustration of the exponents of the security measures C_{1+s} and C_{1-s} (for $s \in [0,1]$) in (55) and (56) respectively for the discrete memoryless multiple source P_{AE} as in Fig. 3. The curves for $C_{1/2}$ and C_1 are identical and they are equal to zero for all rates $R \geq H_1 = 0.4400$ bits per source symbol (cf. Corollary 1).

For the Gallager-type counterparts of the Rényi quantities and $R \geq \hat{R}_1^{\uparrow}$ *(*$\hat{R}_s^{\uparrow}$ *being defined in* (25)*), and any* $s \in [0,1]$*, we also have*

$$\lim_{n\to\infty} -\frac{1}{n} \log \inf_{f_{X_n}} C_{1+s}^{\uparrow}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) = \left| \max_{t\in[s,1]} tH_{1+t}(A|E|P_{AE}) - tR \right|^+, \tag{57}$$

$$\lim_{n\to\infty} -\frac{1}{n} \log \inf_{f_{X_n}} C_{1-s}^{\uparrow}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) = \max_{t\in[0,1]} tH_{1+t}(A|E|P_{AE}) - tR. \tag{58}$$

The infima in (55)*–*(58) *are achieved by any sequence of* ϵ*-almost universal*$_2$ *hash functions* f_{X_n}*.*

This result is proved in Section VII-B. The techniques for the direct parts are somewhat similar to those in [17], [21], [22] using improved versions of Bennett *et al.'s* [7] bound which was based on the Rényi entropy of order 2. However, the non-asymptotic bounds (e.g., Lemma 5) and asymptotic evaluations for the converse parts require new ideas. Different from the direct part, we need to convert the evaluation of $\mathrm{e}^{-sC_{1-s}}$ and $\mathrm{e}^{-\frac{s}{1-s}C_{1-s}^{\uparrow}}$ into information spectrum [47] quantities (involving the conditional entropy random variable) so that is is amenable to asymptotic evaluation. These information spectrum quantities are then evaluated using various large deviation [48] bounds, such as Cramer's theorem. We make several other observations.

First, the exponents of the security indices (namely $-\frac{1}{n}\log C_{1\pm s}$ and $-\frac{1}{n}\log C_{1\pm s}^{\uparrow}$) are non-negative because $C_{1\pm s} \leq \log|f_{X_n}(\mathcal{A}^n)| = O(n)$ and $C_{1\pm s}^{\uparrow} \leq \log|f_{X_n}(\mathcal{A}^n)| = O(n)$ (cf. their definitions in (34) and (36)). The expressions in (56) and (58) are already nonnegative and so we only need to include the $|\cdot|^+$ operation for (55) and (57).

Second, the derivative of the conditional Rényi entropies $\hat{R}_1$ and $\hat{R}_1^{\uparrow}$ are the analogues of the critical rate in error exponent analysis [13], [14]. For the exponents, we have a complete characterization of the exponential rates of decay of both $C_{1\pm s}$ and $C_{1\pm s}^{\uparrow}$ for $s \in [0,1]$ and they are given by optimization of quantities that are related to the conditional Rényi entropy. We observe that the Gallager form results in larger exponents in general as the optimizations in (56) and (58) are larger than their non-Gallager counterparts in (55) and (57) respectively.

Finally, the exponents in (55) and (56) of Theorem 2 are illustrated in Fig. 5. We observe the same behavior for the exponents of the Gallager forms in (57) and (58) since the expressions are the same and so we omit these cases. We note (from the plot and from direct evaluations) that the zero-crossings for the exponents of $C_{1/2}, C_1, C_{3/2}$ and $C_{7/4}$ occur at $H_1, H_1, H_{3/2}$ and $H_{7/4}$ respectively (H_1 being the Shannon entropy). This is in line with Corollary 1. Indeed, the exponent being positive implies that the normalized security measures $\frac{1}{n}C_{1\pm s}$ and $\frac{1}{n}C_{1\pm s}^{\uparrow}$ vanish as blocklength grows. Thus, we conclude that the optimal key generation rates under both the strong and weak secrecy criteria are the same.

V. Second-Order Asymptotics

In the previous sections, the security measures in terms of equivocations and their logarithms were normalized by the blocklength n. In this section, we study different normalizations, e.g., by $\sqrt{n}$. In addition, we examine the effect of changing the size of the hash function M_n from e^{nR} (considered in Sections III and IV) to $\mathrm{e}^{nR+\sqrt{n}L}$, where $L \in \mathbb{R}$ is an arbitrary real number.

A. Basic Definitions

To present our results, we first define the following important quantities.

Definition 2. *Given a discrete joint source $P_{AE} \in \mathcal{P}(\mathcal{A} \times \mathcal{E})$, define the* conditional varentropy *[49] or* conditional source dispersion *[50], [51] to be*

$$V(A|E|P_{AE}) := \sum_{a,e} P_{AE}(a,e)\big(\log P_{A|E}(a|e) + H(A|E|P_{AE})\big)^2. \tag{59}$$

We also define the following variants of the conditional varentropy

$$V_1(A|E|P_{AE}) := \sum_{e} P_E(e)\big(H(A|E|P_{AE}) - H(A|P_{A|E=e})\big)^2 \tag{60}$$

$$V_2(A|E|P_{AE}) := V(A|E|P_{AE}) - V_1(A|E|P_{AE}) \tag{61}$$

$$= \sum_{a,e} P_{AE}(a,e)\big(\log P_{A|E}(a|e) + H(A|P_{A|E=e})\big)^2. \tag{62}$$

One can readily check that $V = V_1 + V_2$ from the definitions. This also follows immediately from the law of total variance. Let

$$\Phi(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \mathrm{e}^{-u^2/2}\,\mathrm{d}u \tag{63}$$

be the cumulative distribution function of the standard Gaussian random variable. With these definitions, we are ready to state our results on the second-order asymptotics for the security measures C_{1+s} and $C_{1+s}^{\uparrow}$ which are simple functions of the equivocation H_{1+s} and $H_{1+s}^{\uparrow}$ respectively. Note that the second-order analysis of C_1 (corresponding to the $s = 0$ case) with no side information (i.e., $E = \emptyset$) was performed in Hayashi's work [16, Theorem 8] in the context of intrinsic randomness based on the relative entropy (Kullback-Leibler divergence) criterion. The other results in Theorems 3 and 4 are novel.

To state our result succinctly, we define the quantities which all depend on s, L and P_{AE} (but we suppress the dependence on the fixed joint distribution P_{AE} for brevity):

$$\Gamma_1(s,L) := -\frac{1}{s}\log\Big(2^{\frac{s}{1-s}} s^{\frac{s}{1-s}}(1-s)\Big) - \frac{1}{s}\log\Phi\bigg(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg), \tag{64}$$

$$\Gamma_2(s,L) := -\frac{1-s}{s}\log\Phi\bigg(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg), \tag{65}$$

$$\Psi_1(s,L) := -\frac{1}{s}\log\Big(2^{s+\frac{s}{1-s}} s^{\frac{s}{1-s}}(1-s)\Big) - \frac{1-s}{s}\log\int_{-\infty}^{\infty}\Phi\bigg(-\frac{L+x}{\sqrt{V_2(A|E|P_{AE})}}\bigg)^{\frac{1}{1-s}} \frac{\mathrm{e}^{-x^2/(2V_1(A|E|P_{AE}))}}{\sqrt{2\pi V_1(A|E|P_{AE})}}\,\mathrm{d}x, \tag{66}$$

$$\Psi_2(s,L) := -\frac{1-s}{s}\log\Phi\bigg(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg). \tag{67}$$

B. Bounds on the Second-Order Asymptotics

Theorem 3 (Second-Order Asymptotics)**.** *Assume that $f_{X_n} : \mathcal{A}^n \to \mathcal{M}_n = \{1, \ldots, M_n\}$ is a random hash function. Consider the following three cases:*

- *Case (A): $\alpha = 1+s$ with $s \in (0,1]$: Suppose that the number of messages $M_n = \mathrm{e}^{nH_{1+s}(A|E|P_{AE})+\sqrt{n}L}$ or $M_n = \mathrm{e}^{nH^{\uparrow}_{1+s}(A|E|P_{AE})+\sqrt{n}L}$. When $L \geq 0$, we have*

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \inf_{f_{X_n}} C_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) = L \tag{68}$$

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \inf_{f_{X_n}} C^{\uparrow}_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) = L. \tag{69}$$

Similarly to Theorem 1, the infima in (68) *and* (69) *are achieved by any sequence of ϵ-almost universal$_2$ hash functions f_{X_n}.*

When $L \leq 0$, we have

$$\lim_{n\to\infty} -\frac{1}{\sqrt{n}} \log \inf_{f_{X_n}} C_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) = -sL \tag{70}$$

$$\lim_{n\to\infty} -\frac{1}{\sqrt{n}} \log \inf_{f_{X_n}} C^{\uparrow}_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) = -sL. \tag{71}$$

Similarly to Theorem 2, the infima in (70) *and* (71) *are achieved by any sequence of ϵ-almost universal$_2$ hash functions f_{X_n}.*

- *Case (B): $\alpha = 1$ (i.e., $s = 0$): Suppose that $M_n = \mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}$ for some $L \in \mathbb{R}$, we have*

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \inf_{f_{X_n}} C_1(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) = \int_{-\infty}^{L/\sqrt{V(A|E|P_{AE})}} \frac{L - \sqrt{V(A|E|P_{AE})}x}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}\, \mathrm{d}x. \tag{72}$$

By (21)*, the same asymptotic behavior also holds true for the Gallager version of the security measure $C^{\uparrow}_1(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})$. Similarly to Theorem 1, the infima in* (72) *is achieved by any sequence of ϵ-almost universal$_2$ hash functions f_{X_n}.*

- *Case (C): $\alpha = 1-s$ with $s \in (0,1]$: Suppose that $M_n = \mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}$ for some $L \in \mathbb{R}$, we have*

$$\begin{aligned}\max\{\Gamma_1(s,L), \Gamma_2(s,L)\} &\leq \liminf_{n\to\infty} \inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \\ &\leq \limsup_{n\to\infty} \inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \leq \frac{\Gamma_2(s,L)}{1-s}.\end{aligned} \tag{73}$$

In addition, for the Gallager-type counterparts, with $M_n = \mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}$ for some $L \in \mathbb{R}$, we also have

$$\begin{aligned}&\max\{\Psi_1(s,L), \Psi_2(s,L)\} \\ &\quad \leq \liminf_{n\to\infty} \inf_{f_{X_n}} C^{\uparrow}_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \\ &\quad \leq \limsup_{n\to\infty} \inf_{f_{X_n}} C^{\uparrow}_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \\ &\quad \leq \Psi_1(s,L) + \frac{1}{s} \log\left(2^{s+\frac{1}{1-s}} s^{\frac{s}{1-s}} (1-s)\right).\end{aligned} \tag{74}$$

The upper bounds in (73) *and* (74) *are achieved by any sequence of ϵ-almost universal$_2$ has functions f_{X_n}.*

This result is proved in Section VII-C. We remark that the converse parts (lower bounds) to (68)–(69) hold for all $s \geq 0$ (and not only being upper bounded by 1) owing to the data processing inequalities in (27)–(28).

C. Remarks on Theorem 3

Observe that in Theorem 3 (Case (A) for instance), the number of compressed symbols M_n satisfies

$$\log M_n = nH_{1+s}(A|E|P_{AE}) + \sqrt{n}L, \qquad \text{or,} \tag{75}$$

$$\log M_n = nH^{\uparrow}_{1+s}(A|E|P_{AE}) + \sqrt{n}L. \tag{76}$$

The leading conditional Rényi entropy terms scaling in n are known as the first-order terms, while the terms scaling as $\sqrt{n}$ are known as the second-order terms. The coefficient L is known as the *second-order coding rate* [16], [52], [53] and the second-order asymptotic characterizations depend on L. Note that even though L is termed as

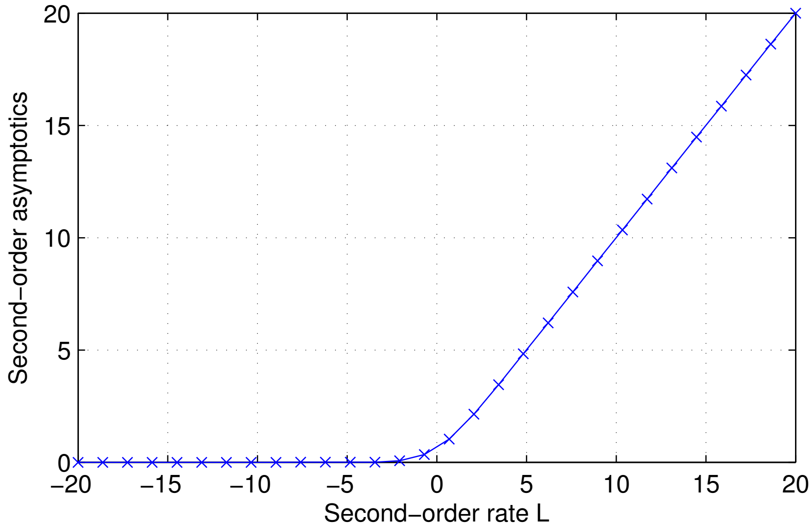

Fig. 6. Illustration of the second-order asymptotics in Case (B) given by the right-hand-side of (72) for the discrete memoryless multiple source P_{AE} as in Fig. 3. It is easy to see that the integral there is non-negative.

the second-order coding *rate*, it may be negative. Observe that the conditional varentropies appear in (72)–(74), which suggests that we evaluate the one-shot bounds using the central limit theorem among other techniques. We have tight results (equalities) for Cases (A) and (B) but unfortunately not for Case (C) where the Rényi parameter $\alpha = 1 - s$ for $s \in (0, 1]$. However, in the limit of the second-order coding rate L being large (either in the positive or negative direction), we can assert that one of the terms in the maxima in the lower bounds of (73) and (74) dominates and matches the upper bound and hence, we have a tight result up to the term in L^2 (Theorem 4). We now comment specifically on each of the cases.

1) For Case (A), the second-order asymptotic behaviors of C_{1+s} and $C^{\uparrow}_{1+s}$ when they are normalized by $\frac{1}{\sqrt{n}}$ are linear in L.
2) The same is true for Case (B) for large positive L because with $V := V(A|E|P_{AE})$,

$$\lim_{L\to\infty} \frac{1}{L} \cdot \int_{-\infty}^{L/\sqrt{V}} \frac{L - \sqrt{V}x}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}\, \mathrm{d}x = \lim_{L\to\infty} \left\{ \Phi\Big(\frac{L}{\sqrt{V}}\Big) - \frac{\sqrt{V}}{L} \int_{-\infty}^{L/\sqrt{V}} \frac{x\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}}\, \mathrm{d}x \right\} = 1. \tag{77}$$

In contrast, when $L \to -\infty$ in Case (B), the limit is zero. The second-order asymptotics in Case (B) in (72) is shown in Fig. 6 and is obtained via numerical integration to approximate the integral. The limit in (72) is monotonically increasing in L. This is intuitive because as L increases, there is potentially more leakage to E^n and less uniformity on the (larger) support $\{1, \ldots, \mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}\}$.
3) For Case (C) there is no normalization by $\frac{1}{\sqrt{n}}$ and we only have bounds. However, for large $|L|$, we will see from Theorem 4 that the second-order asymptotic behavior is quadratic in L. The bounds on the second-order asymptotics in the two parts (conditional Rényi entropy and its Gallager version) of Case (C) in (73) and (74) are shown in Figs. 7 and 8 respectively.

We conclude that in the second-order asymptotic regime where the number of compressed symbols satisfies (75)–(76), there are distinct differences between the three regimes of the Rényi parameter $\alpha \in [0, 1)$, $\alpha = 1$ and $\alpha \in (1, 2]$.

D. Approximations for Large Second-Order Coding Rates

Since for Case (C) we only have bounds, we now examine the behavior of the bounds in the limit of large $|L|$ for which we can show tight results up to the quadratic terms.

Theorem 4 (Large Second-Order Rates)**.** *Assume that $f_{X_n} : \mathcal{A}^n \to \mathcal{M}_n = \{1, \ldots, M_n\}$ is a random hash function. For Case (C) in Theorem 3 (Rényi parameter $\alpha = 1 - s$ where $s \in (0, 1]$), we have the following asymptotic results*

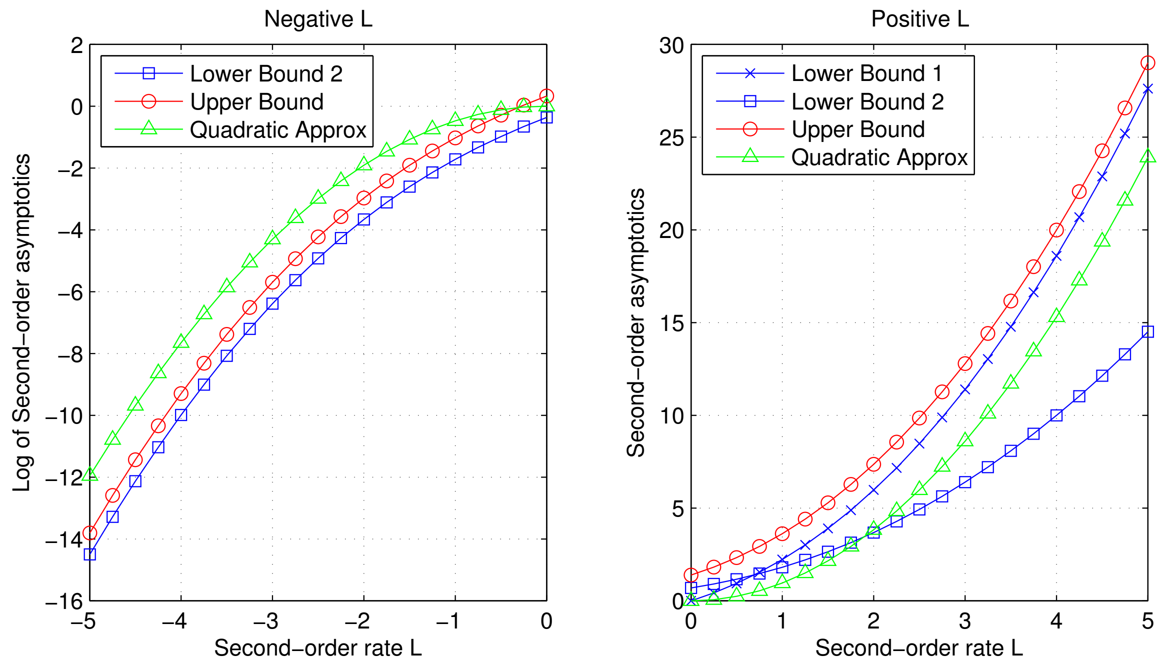

Fig. 7. Illustration of the bounds on the second-order asymptotics of $C_{1/2}(f_{X_n}(A^n)|E^nX_n|P_{AE}^n \times P_{X_n})$ (i.e., $s = 1/2$) in Case (C) given by the left- and right-hand-sides of (73) for the same source. Note that the figure on the left is plotted in log scale (corresponding to (80)) while the figure on the right is plotted in linear scale. For $L \le 0$, lower bound 1 in (73) is negative (lower bound 2 dominates) so is not shown in the left plot. Observe the quadratic behaviors; this is corroborated by Theorem 4. The quadratic approximations in (78) and (80) (without the $O(\log L)$ terms) are also plotted. Observe that there is a constant offset between the quadratic and the bounds as we do not determine the $O(\log L)$ terms in Theorem 4 exactly.

as $L \to \infty$:

$$\lim_{n\to\infty} \inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^nX_n|P_{AE}^n \times P_{X_n}) = \frac{L^2}{2sV(A|E|P_{AE})} + O(\log L), \tag{78}$$

$$\lim_{n\to\infty} \inf_{f_{X_n}} C_{1-s}^{\uparrow}(f_{X_n}(A^n)|E^nX_n|P_{AE}^n \times P_{X_n}) = \frac{1-s}{2s} \cdot \frac{L^2}{V_1(A|E|P_{AE}) + V_2(A|E|P_{AE})(1-s)} + O(\log L). \tag{79}$$

Furthermore, we have the following asymptotic results as $L \to -\infty$:

$$\lim_{n\to\infty} \log \inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^nX_n|P_{AE}^n \times P_{X_n}) = -\frac{L^2}{2V(A|E|P_{AE})} + O(\log |L|), \tag{80}$$

$$\lim_{n\to\infty} \log \inf_{f_{X_n}} C_{1-s}^{\uparrow}(f_{X_n}(A^n)|E^nX_n|P_{AE}^n \times P_{X_n}) = -\frac{L^2}{2V(A|E|P_{AE})} + O(\log |L|). \tag{81}$$

The infima in (78)–(81) *are achieved by any sequence of ϵ-almost universal$_2$ hash functions f_{X_n}.*

The proof of Theorem 4 can be found in Section V-E.

The results in Theorem 4 are somewhat analogous and similar those in the study of the *moderate-deviations* asymptotics in information theory [54]–[58]. Note the difference between the results in (78)–(79) ($L \to \infty$) versus (80)–(81) ($L \to -\infty$). The former pair of results resembles the equivocation results presented in Section III since the effective rate is $(L/\sqrt{n})$-*higher* than the conditional Rényi entropy and there is no logarithm preceding C_{1-s} and $C_{1-s}^{\uparrow}$. The latter pair of results resembles the exponent results of Section IV since the effective rate is $(|L|/\sqrt{n})$-*lower* than the conditional Rényi entropy and there is a logarithm preceding C_{1-s} and $C_{1-s}^{\uparrow}$. So the results presented in Theorem 4 are natural in view of the equivocation result in Theorem 1 and the exponent result in Theorem 2.

E. Proof of Theorem 4

In this section, we present the proof of Theorem 4. Since there are four statements in (78)–(81), we partition the proof into four distinct subsections.

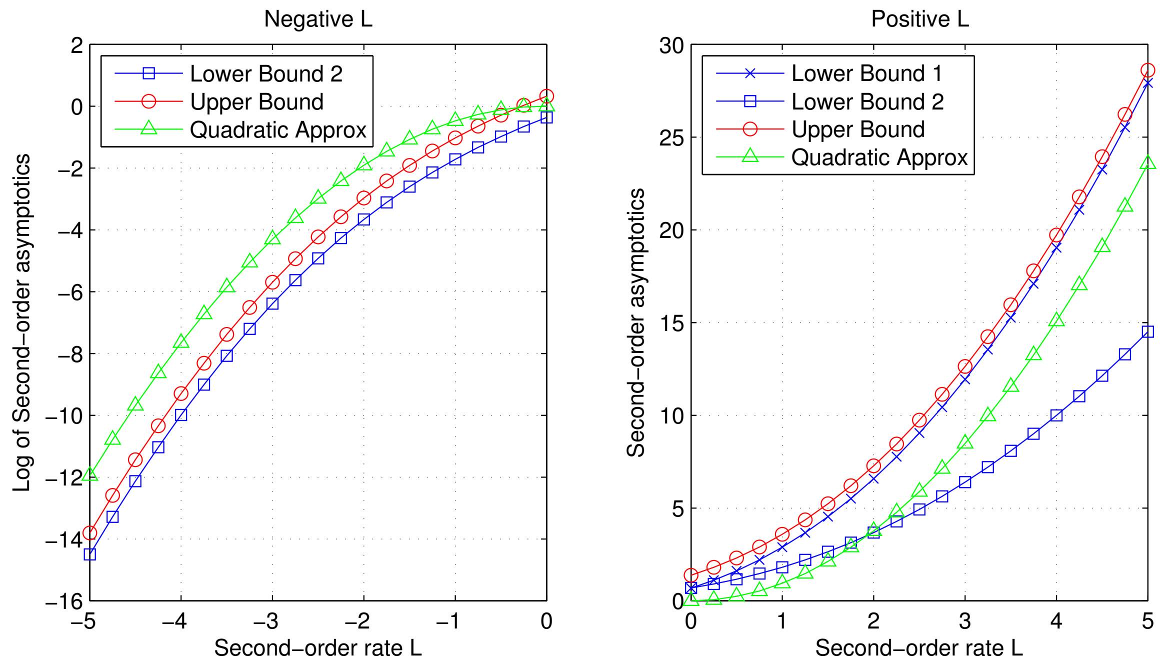

Fig. 8. Illustration of the bounds on the second-order asymptotics of $C^{\uparrow}_{1/2}(f_{X_n}(A^n)|E^nX_n|P^n_{AE}\times P_{X_n})$ (i.e., $s=1/2$) in Case (C) given by the left- and right-hand-sides of (74) for the same source. For $L\le 0$, lower bound 1 in (74) is negative (lower bound 2 dominates) so is not shown in the left plot. Observe the quadratic behaviors–this is corroborated by (79) and (81) in Theorem 4. These plots are obtained by using numerical integration to calculate the integral in $\Psi_1(s,L)$ in (74).

a) Proof of (78)*:* When $L\to\infty$, the term $\Gamma_1(s,L)$ in (64) behaves as $-\frac{1}{s}\log\Phi\big(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\big)+O(1)$, and attains the maximum in the lower bound in (73) because $\Gamma_2(s,L)$ has the additional factor $1-s$, which is smaller than 1. Also see the right plot of Fig. 7. Thus, in this limiting regime, the lower bound matches the upper bound in (73), namely $-\frac{1}{s}\log\Phi\big(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\big)$, up to a constant term, i.e.,

$$\lim_{n\to\infty}\inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^nX_n|P^n_{AE}\times P_{X_n}) = -\frac{1}{s}\log\Phi\bigg(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg)+O(1) \tag{82}$$

where $O(1)$ denotes a term bounded in L (but dependent on s). Now by employing the asymptotic equality

$$\Phi(-t)=1-\Phi(t)\sim\frac{\mathrm{e}^{-t^2/2}}{\sqrt{2\pi}t},\qquad \text{as } t\to\infty, \tag{83}$$

we obtain from (82) that

$$\lim_{n\to\infty}\inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^nX_n|P^n_{AE}\times P_{X_n}) = \frac{L^2}{2sV(A|E|P_{AE})}+O(\log L), \tag{84}$$

which proves (78).

b) Proof of (79)*:* When $L\to\infty$, the term $\Psi_1(s,L)$ dominates the maximum in the lower bound in (74) because $V_2\le V$ and thus the integrands in $\Psi_1(s,L)$, namely $\Phi\big(-\frac{L+x}{\sqrt{V_2(A|E|P_{AE})}}\big)^{\frac{1}{1-s}}\frac{\mathrm{e}^{-x^2/(2V_1(A|E|P_{AE}))}}{\sqrt{2\pi V_1(A|E|P_{AE})}}$, are not smaller than $\Phi\big(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\big)$ in $\Psi_2(s,L)$. See right plot of Fig. 8. We can then find the x that dominates the integral in $\Psi_1(s,L)$. We denote this by x^*. Since L is large, by (83),

$$\log\bigg[\Phi\bigg(-\frac{L+x}{\sqrt{V_2}}\bigg)^{\frac{1}{1-s}}\mathrm{e}^{-x^2/(2V_1)}\bigg]-\bigg[-\frac{1}{2(1-s)}\bigg(\frac{L+x}{\sqrt{V_2}}\bigg)^2-\frac{x^2}{2V_1}\bigg]\to 0,\quad \text{as } L\to\infty. \tag{85}$$

Differentiating the quadratic, we obtain

$$x^*=-\frac{LV_1(A|E|P_{AE})}{V_1(A|E|P_{AE})+(1-s)V_2(A|E|P_{AE})}. \tag{86}$$

The exponential term $\mathrm{e}^{-(x^*)^2/(2V_1(A|E|P_{AE}))}$ controls the behavior of the integral in $\Psi_1(s,L)$ and substituting (86) into this exponential term yields (79).

c) Proof of (80)*:* Now we assume that $L \to -\infty$. In this case, we find that the term $\Gamma_2(s,L)$ attains the maximum in the lower bound in (73) because $\Gamma_1(s,L)$ is negative due to the constant negative term. Also see the left plot of Fig. 7. In this case, taking the logarithm, we have

$$\begin{aligned}
&\lim_{n\to\infty} \log \inf_{f_{X_n}} C_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \\
&= \log\left[-\log\left(1-\Phi\left(\frac{L}{\sqrt{V(A|E|P_{AE})}}\right)\right)\right] + O(1) && (87)\\
&= \log\left[\Phi\left(\frac{L}{\sqrt{V(A|E|P_{AE})}}\right)\right] + O(\log|L|) && (88)\\
&= -\frac{L^2}{2V(A|E|P_{AE})} + O(\log|L|) && (89)
\end{aligned}$$

where in (87), $\log(\frac{1-s}{s})$ and $\log(\frac{1}{s})$ can be regarded as $O(1)$ when $L \to -\infty$, in (88), we used the fact that $\log(1-t) = -t + O(t^2)$ when $t \downarrow 0$, and finally in (89), we used (83). This proves (80).

d) Proof of (81)*:* In the other direction, when $L \to -\infty$, we claim that the term $\Psi_2(s,L)$ attains the maximum. This is shown as follows: First, we find that

$$\left[1-\Phi\left(\frac{L+x}{\sqrt{V_2(A|E|P_{AE})}}\right)\right]^{\frac{1}{1-s}} \geq 1 - \frac{1}{1-s}\Phi\left(\frac{L+x}{\sqrt{V_2(A|E|P_{AE})}}\right). \tag{90}$$

This is because $a \mapsto a^{\frac{1}{1-s}}$ is convex and so the linear approximation underestimates the function. This means that

$$\begin{aligned}
&\log \int_{-\infty}^{\infty} \Phi\left(-\frac{L+x}{\sqrt{V_2(A|E|P_{AE})}}\right)^{\frac{1}{1-s}} \frac{\mathrm{e}^{-x^2/(2V_1(A|E|P_{AE}))}}{\sqrt{2\pi V_1(A|E|P_{AE})}}\,\mathrm{d}x \\
&\geq \log \int_{-\infty}^{\infty} \left[1-\frac{1}{1-s}\Phi\left(\frac{L+x}{\sqrt{V_2(A|E|P_{AE})}}\right)\right] \frac{\mathrm{e}^{-x^2/(2V_1(A|E|P_{AE}))}}{\sqrt{2\pi V_1(A|E|P_{AE})}}\,\mathrm{d}x && (91)\\
&= \log\left[1-\frac{1}{1-s}\Phi\left(\frac{L}{\sqrt{V(A|E|P_{AE})}}\right)\right] && (92)\\
&= -\frac{1}{1-s}\Phi\left(\frac{L}{\sqrt{V(A|E|P_{AE})}}\right) + O\left(\mathrm{e}^{-L^4/(4V(A|E|P_{AE})^2)}\right) && (93)\\
&= -\frac{1}{1-s}\log\left[1-\Phi\left(\frac{L}{\sqrt{V(A|E|P_{AE})}}\right)\right] + O\left(\mathrm{e}^{-L^4/(4V(A|E|P_{AE})^2)}\right) && (94)\\
&= -\frac{1}{1-s}\log\Phi\left(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\right) + O\left(\mathrm{e}^{-L^4/(4V(A|E|P_{AE})^2)}\right), && (95)
\end{aligned}$$

where (92) follows because the convolution of two independent zero-mean Gaussians is a Gaussian where the variances add and we also note that $V = V_1 + V_2$ per (61). This argument was also used in the second-order analysis of channels with state [59, Lemma 18]. Inequalities (93) and (94) follow from the fact that $\log(1-x) = -x + O(x^2)$ as $x \downarrow 0$ (note that $L \to -\infty$ so the term $\Phi(L/\sqrt{V})$ tends to zero). Hence, (95) and the definitions of $\Psi_2(s,L)$ and $\Psi_1(s,L)$ (in (66) and (67) resp.) imply that $\Psi_2(s,L)$ asymptotically dominates $\Psi_1(s,L)$ as $L \to -\infty$. Also

see left plot of Fig. 8. By a similar calculation as in (87)–(89), we obtain the lower bound to (81) as follows:

$$\liminf_{n\to\infty} \log \inf_{f_{X_n}} C^{\uparrow}_{1-s}(f_{X_n}(A^n)|E^n X_n|P^n_{AE} \times P_{X_n})$$

$$\geq \log\left[-\frac{1-s}{s}\log\Phi\bigg(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg)\right] + O(1) \tag{96}$$

$$= \log\left[-\frac{1-s}{s}\left(-\Phi\bigg(\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg)\right)\right] + O(\log|L|)) \tag{97}$$

$$= \log\left[\Phi\bigg(\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg)\right] + O(\log|L|) \tag{98}$$

$$= -\frac{L^2}{2V(A|E|P_{AE})} + O(\log|L|), \tag{99}$$

where in (96), we used the above observation that $\Psi_2(s,L) = \Omega(\Psi_1(s,L))$ as $L \to -\infty$, in (97), we wrote $\Phi(-L/\sqrt{V}) = 1 - \Phi(L/\sqrt{V})$ and used the fact that $\log(1-x) = -x + O(x^2)$ as $x \downarrow 0$, in (98) we used the fact that $\log[(1-s)/s] = O(1)$, and finally in (99) we used the approximation in (83). To show that the upper bound in (74) matches the lower bound given by (99) (when $L \to -\infty$), we use (90) and steps similar to those in (91)–(95) to assert that

$$-\frac{1-s}{s}\log\int_{-\infty}^{\infty}\Phi\bigg(-\frac{L+x}{\sqrt{V_2(A|E|P_{AE})}}\bigg)^{\frac{1}{1-s}}\frac{\mathrm{e}^{-x^2/(2V_1(A|E|P_{AE}))}}{\sqrt{2\pi V_1(A|E|P_{AE})}}\,\mathrm{d}x$$

$$\leq -\frac{1-s}{s}\log\left[1 - \frac{1}{1-s}\Phi\bigg(\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg)\right] \tag{100}$$

$$= \frac{1}{s}\Phi\bigg(\frac{L}{\sqrt{V(A|E|P_{AE})}}\bigg) + O\left(\mathrm{e}^{-L^4/(4V(A|E|P_{AE})^2)}\right), \tag{101}$$

where in (101), we again used the fact that $\log(1-x) = -x + O(x^2)$ for $x \downarrow 0$. Now taking the logarithm and the limit as $L \to -\infty$, we match the lower bound in (99) completing the proof of (81).

VI. One-Shot Bounds

To prove Theorems 1, 2 and 3, we leverage the following one-shot (i.e., blocklength n equal to 1) bounds. The proofs of these one-shot bounds are rather technical and hence we provide them in the appendices.

A. *One-Shot Bounds for the Direct Parts*

For the direct parts of the equivocation results, we evaluate the following one-shot bounds. The first two bounds in (102) and (103) can be considered as generalizations of the bounds by Hayashi in [17] where $\epsilon = 1$.

Lemma 1. *For an ensemble of ϵ-almost universal$_2$ hash functions $f_X : \mathcal{A} \to \mathcal{M} = \{1,\ldots,M\}$, we have for $s \in [0,1]$,*

$$\mathrm{e}^{sC_{1+s}(f_X(A)|EX|P_{AE}\times P_X)} \leq \epsilon^s + M^s \mathrm{e}^{-sH_{1+s}(A|E|P_{AE})}, \tag{102}$$

$$\mathrm{e}^{\frac{s}{1+s}C^{\uparrow}_{1+s}(f_X(A)|EX|P_{AE}\times P_X)} \leq \epsilon^{\frac{s}{1+s}} + M^{\frac{s}{1+s}}\mathrm{e}^{-\frac{s}{1+s}H^{\uparrow}_{1+s}(A|E|P_{AE})}. \tag{103}$$

In the other direction with $s \in [0,1]$,

$$\mathrm{e}^{-sC_{1-s}(f_X(A)|EX|P_{AE}\times P_X))}$$

$$\geq 2^{-s}\sum_{(a,e):P_{A|E}(a|e)\geq\frac{\epsilon}{M}} P_{AE}(a,e)P_{A|E}(a|e)^{-s}M^{-s} + 2^{-s}\sum_{(a,e):P_{A|E}(a|e)<\frac{\epsilon}{M}} P_{AE}(a,e)\epsilon^{-s}, \tag{104}$$

$$\mathrm{e}^{-\frac{s}{1-s}C^{\uparrow}_{1-s}(f_X(A)|EX|P_{AE}\times P_X))}$$

$$\geq \frac{1}{2M^{\frac{s}{1-s}}}\sum_e P_E(e)\bigg(\sum_{a:P_{A|E}(a|e)\geq\frac{\epsilon}{M}} P_{A|E}(a|e)^{1-s}\bigg)^{\frac{1}{1-s}} + (2\epsilon)^{-\frac{s}{1-s}}\sum_e P_E(e)\bigg(\sum_{a:P_{A|E}(a|e)<\frac{\epsilon}{M}} P_{A|E}(a|e)\bigg)^{\frac{1}{1-s}}. \tag{105}$$

For the direct parts of the exponents results, we evaluate the following one-shot bound.

Lemma 2. *For an ensemble of universal$_2$ hash functions $f_X : \mathcal{A} \to \mathcal{M} = \{1, \dots, M\}$, we have for any $s \in [0, 1]$,*

$$\mathrm{e}^{\frac{s}{1+s} C^{\uparrow}_{1+s}(f_X(A)|EX|P_{AE}\times P_X)} \le 1 + \frac{1}{1+s} M^s \mathrm{e}^{-sH_{1+s}(A|E|P_{AE})}. \tag{106}$$

For the direct parts of the second-order results, we evaluate the following one-shot bound.

Lemma 3. *For an ensemble of an ϵ-almost universal$_2$ hash functions $f_X : \mathcal{A} \to \mathcal{M} = \{1, \dots, M\}$, we have for any $s \in [0, 1]$ and $c > 0$,*

$$\mathrm{e}^{-sC_{1-s}(f_X(A)|EX|P_{AE}\times P_X)} \ge P_{AE}\Big\{(a, e) : P_{A|E}(a|e) \le \frac{c}{M}\Big\}\Big(\frac{1}{c+\epsilon}\Big)^s, \tag{107}$$

$$\mathrm{e}^{-\frac{s}{1-s} C^{\uparrow}_{1-s}(f_X(A)|EX|P_{AE}\times P_X)} \ge \Big(\frac{1}{c+\epsilon}\Big)^{\frac{s}{1-s}} \sum_e P_E(e) \Bigg(\sum_{a: P_{A|E}(a|e) \le \frac{c}{M}} P_{A|E}(a|e) \Bigg)^{\frac{1}{1-s}}. \tag{108}$$

B. One-Shot Bounds for the Converse Parts

For the converse parts of the equivocation results, we evaluate the following one-shot bounds.

Lemma 4. *Fix $c > 1$ and $s \ge 0$. Any hash function $f : \mathcal{A} \to \mathcal{M} = \{1, \dots, M\}$ satisfies*

$$\begin{aligned}\mathrm{e}^{-sC_{1-s}(f(A)|E|P_{AE})} \le c^{-s} \sum_e P_E(e) \sum_{a: P_{A|E}(a|e) \ge \frac{c}{M}} P_{A|E}(a|e)^{1-s} M^{-s} \\ + 2^{\frac{s}{1-s}} s^{\frac{s}{1-s}} (1-s) P_{AE}\Big\{(a, e) : P_{A|E}(a|e) \le \frac{c}{M}\Big\}.\end{aligned} \tag{109}$$

For the Gallager-type counterpart,

$$\begin{aligned}\mathrm{e}^{-\frac{s}{1-s} C^{\uparrow}_{1-s}(f(A)|E|P_{AE})} \le 2^{\frac{s}{1-s}} \sum_e P_E(e) \Bigg[\Bigg(c^{-s} \sum_{a: P_{A|E}(a|e) \ge \frac{c}{M}} P_{A|E}(a|e)^{1-s} M^{-s} \Bigg)^{\frac{1}{1-s}} \\ + \Big(2^{\frac{s}{1-s}} s^{\frac{s}{1-s}} (1-s) P_{A|E=e}\Big\{a : P_{A|E}(a|e) < \frac{c}{M}\Big\}\Big)^{\frac{1}{1-s}} \Bigg].\end{aligned} \tag{110}$$

For the converse parts of the exponents results and the second-order results for the Rényi parameter being $1 + s$ (with s nonnegative), we evaluate the following one-shot bounds.

Lemma 5. *Fix $c > 1$ and $s \in [0, 1]$. Any hash function $f : \mathcal{A} \to \mathcal{M} = \{1, \dots, M\}$ satisfies*

$$\mathrm{e}^{-sC_{1-s}(f(A)|E|P_{AE})} \le \sum_{(a,e): P_{A|E}(a|e) \ge \frac{c}{M}} P_E(e) P_{A|E}(a|e)^{1-s} M^{-s} + \sum_e P_E(e) P_{A|E=e}\Big\{a : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1-s} \tag{111}$$

$$\le P_{A,E}\Big\{(a, e) : P_{A|E}(a|e) \ge \frac{c}{M}\Big\} c^{-s} + P_{A,E}\Big\{(a, e) : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1-s}. \tag{112}$$

For the Gallager-type counterpart, for $s \in [0, 1)$,

$$\begin{aligned}\mathrm{e}^{-\frac{s}{1-s} C^{\uparrow}_{1-s}(f(A)|E|P_{AE})} \le \sum_e P_E(e) \Bigg[P_{A|E=e}\Big\{a : P_{A|E}(a|e) \ge \frac{c}{M}\Big\} c^{-s} \\ + P_{A|E=e}\Big\{a : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1-s} \Bigg]^{\frac{1}{1-s}}.\end{aligned} \tag{113}$$

In the other direction, for $s \in [0, 1]$, we have

$$\begin{aligned}\mathrm{e}^{sC_{1+s}(f(A)|E|P_{AE})} \ge \sum_{(a,e): P_{A|E}(a|e) \ge \frac{c}{M}} P_E(e) P_{A|E}(a|e)^{1+s} M^s \\ + \sum_e P_E(e) P_{A|E=e}\Big\{(a, e) : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1+s}\end{aligned} \tag{114}$$

$$\ge P_{AE}\Big\{(a, e) : P_{A|E}(a|e) \ge \frac{c}{M}\Big\} c^s + P_{AE}\Big\{(a, e) : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1+s}. \tag{115}$$

For the Gallager-type counterpart, for $s \in [0,1]$, we have

$$\mathrm{e}^{\frac{s}{1+s} C^{\uparrow}_{1+s}(f(A)|E|P_{AE})} \geq \sum_{e} P_E(e) \bigg[P_{A|E=E}\Big\{ a : P_{A|E}(a|e) \geq \frac{c}{M} \Big\} c^s + P_{A|E=E}\Big\{ a : P_{A|E}(a|e) < \frac{c}{M} \Big\}^{1+s} \bigg]^{\frac{1}{1+s}}. \tag{116}$$

For the converse parts of the second-order results for the Rényi parameter being $1-s$ (with s nonnegative), we need the following one-shot bound as well as (110) and (113) although the converse parts of the second-order results with Rényi parameter being $1+s$ require (114) and (116).

Lemma 6. *Fix $c > 1$ and $s \in [0,1]$. Any hash function $f : \mathcal{A} \to \mathcal{M} = \{1, \ldots, M\}$ satisfies*

$$\mathrm{e}^{-sC_{1-s}(f(A)|E|P_{AE})} \leq c^{-s} P_{AE}\Big\{ (a,e) : P_{A|E}(a|e) \geq \frac{c}{M} \Big\} + 2^{\frac{s}{1-s}} s^{\frac{s}{1-s}} P_{AE}\Big\{ (a,e) : P_{A|E}(a|e) \leq \frac{c}{M} \Big\}. \tag{117}$$

VII. Proofs of the Asymptotic Results

In this section, we prove the asymptotic results in Theorems 1, 2, and 3.

Notation: Throughout, we let $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathcal{A}^n$ and $\mathbf{e} = (e_1, e_2, \ldots, e_n) \in \mathcal{E}^n$ denote deterministic length-n strings. We also let $A^n = (A_1, A_2, \ldots, A_n)$ and $E^n = (E_1, E_2, \ldots, E_n)$ denote random vectors of length n. We adopt the exponential equality notation: $a_n \doteq b_n$ if and only if $\lim_{n\to\infty} \frac{1}{n} \log \frac{a_n}{b_n} = 0$.

Given a random variable X with distribution (probability mass function) P, we denote the expectation of a function of the random variable $g(X)$ by $\mathbb{E}[g(X)] = \sum_x P(x) g(x)$. If we want to make the dependence of the expectation on X or P explicit, we write $\mathbb{E}_X[g(X)]$ or $\mathbb{E}_P[g(X)]$. The same comment applies to the variance operator which we denote interchangeably as $\mathsf{Var}[g(X)]$, $\mathsf{Var}_X[g(X)]$ or $\mathsf{Var}_P[g(X)]$.

A. *Proof of Theorem 1*

1) Direct Parts: We first prove the direct parts (upper bounds).

a) Proof of the upper bound of (47)*:* The bound in (102) implies that

$$\begin{aligned} & C_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \\ & \leq \frac{1}{s} \log \big(\epsilon^s + M_n^s \mathrm{e}^{-sH_{1+s}(A^n|E^n|P^n_{AE})} \big) && (118) \\ & = \frac{1}{s} \log \big(\epsilon^s + M_n^s \mathrm{e}^{-nsH_{1+s}(A|E|P_{AE})} \big). && (119) \end{aligned}$$

For ϵ being a constant, this achieves the upper bound of (47) upon normalizing by n and taking the $\limsup$.

b) Proof of the upper bound of (48)*:* The bound in (103) implies that

$$\begin{aligned} & C^{\uparrow}_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \\ & \leq \frac{1+s}{s} \log \big(\epsilon^{\frac{s}{1+s}} + M_n^{\frac{s}{1+s}} \mathrm{e}^{-\frac{s}{1+s} H_{1+s}(A^n|E^n|P^n_{AE})} \big) && (120) \\ & = \frac{1+s}{s} \log \big(\epsilon^{\frac{s}{1+s}} + M_n^{\frac{s}{1+s}} \mathrm{e}^{-n\frac{s}{1+s} H_{1+s}(A|E|P_{AE})} \big). && (121) \end{aligned}$$

This leads to the upper bound of (48) for constant ϵ upon normalizing by n and taking the $\limsup$.

c) Proof of the upper bound of (49)*:* To obtain (49), we employ Cramér's theorem [48] on the sequence of random variables $-\log P^n_{A|E}(A^n|E^n) = \sum_{i=1}^n - \log P_{A|E}(A_i|E_i)$ under the product joint distribution P^n_{AE}. It is easy to see by using exponential tail bounds that

$$\begin{aligned} & \lim_{n\to\infty} -\frac{1}{n} \log \sum_{(\mathbf{a},\mathbf{e}) : P^n_{A|E}(\mathbf{a}|\mathbf{e}) < \mathrm{e}^{-nR}} P^n_{AE}(\mathbf{a},\mathbf{e}) \\ & = \lim_{n\to\infty} -\frac{1}{n} \log P^n_{AE} \left\{ (\mathbf{a},\mathbf{e}) : \frac{1}{n} \sum_{i=1}^n \log P_{A|E}(a_i|e_i) \leq -R \right\} && (122) \\ & = \max_{t \geq 0} t(R - H_{1-t}(A|E|P_{AE})). && (123) \end{aligned}$$

Note that the cumulant generating function of the random variable $-\log P_{A|E}(A|E)$ under the joint distribution P_{AE} can be expressed in terms of the conditional Rényi entropy as

$$\log \mathbb{E}_{P_{AE}}\left[\mathrm{e}^{t(-\log P_{A|E}(A|E))}\right] = tH_{1-t}(A|E|P_{AE}), \tag{124}$$

explaining the presence of this term in (123). We again apply (a generalized version of) Cramer's theorem[4] to the sequence of random variables $\log P^n_{A|E}(A^n|E^n)$ under the sub-distribution (non-negative product measure) $P^n_{AE}(\mathbf{a},\mathbf{e})(P^{-s}_{A|E})^n(\mathbf{a}|\mathbf{e})$ and event $\{(\mathbf{a},\mathbf{e}) : \log P^n_{A|E}(\mathbf{a}|\mathbf{e}) \geq -nR\}$. Note that the cumulant generating function in this case is

$$\begin{aligned}
\tau_s(t) &:= \log \sum_{a,e} P_{AE}(a,e) P^{-s}_{A|E}(a|e) \exp\left(t \log P_{A|E}(a|e)\right) && (125)\\
&= (s-t) H_{1-(s-t)}(A|E|P_{AE}), && (126)
\end{aligned}$$

and by direct differentiation, we also have that

$$\tau_s'(0) = -\hat{R}_{-s} \tag{127}$$

where $\hat{R}_s$ is defined in (24) (cf. $\hat{R}_{-s}$ is presented in a different form in (51)). Thus, by Cramér's theorem,

$$\begin{aligned}
\sum_{(\mathbf{a},\mathbf{e}):P^n_{A|E}(\mathbf{a}|\mathbf{e})\geq \mathrm{e}^{-nR}} P^n_{AE}(\mathbf{a},\mathbf{e}) P^n_{A|E}(\mathbf{a}|\mathbf{e})^{-s} &\doteq \frac{\mathbb{E}[\mathrm{e}^{t\log P_{A|E}(A|E)}]}{\mathrm{e}^{-tnR}} && (128)\\
&= \exp\left[-n\left(-tR - \log \sum_{a,e} P_{AE}(a,e) P_{A|E}(a|e)^s \mathrm{e}^{t\log P_{A|E}(A|E)}\right)\right] && (129)\\
&= \exp\left[-n\left(-tR + (t-s) H_{1+(t-s)}(A|E|P_{AE})\right)\right], && (130)
\end{aligned}$$

where in (128), the "expectation" $\mathbb{E}$ is taken with respect to the non-negative measure $(a,e) \mapsto P_{AE}(a,e)P_{A|E}(a|e)$. Since $t \geq 0$ is arbitrary,

$$\begin{aligned}
&-\frac{1}{n}\log \sum_{(\mathbf{a},\mathbf{e}):P^n_{A|E}(\mathbf{a}|\mathbf{e})\geq \mathrm{e}^{-nR}} P^n_{AE}(\mathbf{a},\mathbf{e}) P^n_{A|E}(\mathbf{a}|\mathbf{e})^{-s}\\
&\qquad = \max_{t\geq 0}\left\{ -tR + (t-s) H_{1+(t-s)}(A|E|P_{AE})\right\} && (131)
\end{aligned}$$

For the case where $R \geq \hat{R}_{-s}$, the constraint in the optimization above is active, i.e., $t^* = 0$ because the function $s \mapsto \hat{R}_{-s}$ is monotonically non-decreasing as described in Section II-A. Conversely, when $R \leq \hat{R}_{-s}$, the constraint is inactive, i.e., the maximum is realized with $R = \hat{R}_{-(s-t)}$. Thus, we obtain

$$\begin{aligned}
&\lim_{n\to\infty} -\frac{1}{n}\log \sum_{(\mathbf{a},\mathbf{e}):P^n_{A|E}(\mathbf{a}|\mathbf{e})\geq \mathrm{e}^{-nR}} P^n_{AE}(\mathbf{a},\mathbf{e}) P^n_{A|E}(\mathbf{a}|\mathbf{e})^{-s} \mathrm{e}^{-snR}\\
&\qquad = \begin{cases} s(R - H_{1-s}(A|E|P_{AE})) & \text{if } R \geq \hat{R}_{-s} \\ \max_{t'\leq s} t'(R - H_{1-t'}(A|E|P_{AE})) & \text{if } R \leq \hat{R}_{-s} \end{cases}, && (132)
\end{aligned}$$

where the second clause follows by the substitution $t' = s - t$. Now with these preparations, we can employ the one-shot bound in (104) with $\epsilon = 1$ to prove the direct part of (49) as follows: Since (132) is not greater than

[4]The standard Cramér's theorem [48, Section 2.2] (or Sanov's theorem [48, Section 2.1]) is a large-deviations result concerning the exponent of $P^n(\mathcal{B})$ where P is a *probability* measure and $\mathcal{B}$ is an event in the sample space Ω. If P is not necessarily a probability measure but a finite non-negative measure (as it is in our applications), say μ, Cramér's theorem clearly also applies by defining the new *probability* measure $\mathcal{B} \mapsto \widetilde{P}(\mathcal{B}) := \mu(\mathcal{B})/\mu(\Omega)$.

(123), the former dominates in the exponent and we obtain

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} C_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \\
&= -\frac{1}{s} \liminf_{n\to\infty} \frac{1}{n} \log \left[\mathrm{e}^{-sC_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})} \right] \qquad (133)\\
&\leq -\frac{1}{s} \liminf_{n\to\infty} \frac{1}{n} \log \Bigg[2^{-s} \sum_{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) < \mathrm{e}^{-nR}} P_{AE}^n(\mathbf{a},\mathbf{e}) \\
&\qquad + 2^{-s} \sum_{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \mathrm{e}^{-nR}} P_{AE}^n(\mathbf{a},\mathbf{e}) P_{A|E}^n(\mathbf{a}|\mathbf{e})^{-s} \mathrm{e}^{-snR} \Bigg], \qquad (134)
\end{aligned}$$

where (134) follows from (104). Now we combine the asymptotic results in (123) and (132) to evaluate the asymptotic behavior of (134). In particular, we take into consideration the scaling factor $\frac{1}{s}$. We also note that the domain of maximization of t in (123) and t' in the second clause of (132) are $[0, \infty)$ and $(\infty, s]$ respectively. So the intersection of these domains is $[0, s]$ and the eventual max should thus be taken over $[0, s]$. Uniting these observations, we obtain that the upper bound

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} C_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \\
&\quad \leq \begin{cases} R - H_{1-s}(A|E|P_{AE}) & \text{if } R \geq \hat{R}_{-s} \\ \max_{t\in[0,s]} \frac{t}{s}(R - H_{1-t}(A|E|P_{AE})) & \text{if } R \leq \hat{R}_{-s} \end{cases}. \qquad (135)
\end{aligned}$$

d) Proof of the upper bound of (50)*:* The upper bound of (50) proceeds in an analogous manner. It proceeds in five distinct steps, each detailed in the following five paragraphs.

In Step 1, we manipulate the one-shot bound in (105) with $\epsilon = 1$ as follows:

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} C_{1-s}^{\uparrow}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \\
&= -\frac{1-s}{s} \liminf_{n\to\infty} \frac{1}{n} \log \left[\mathrm{e}^{-\frac{s}{1-s} C_{1-s}^{\uparrow}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})} \right] \qquad (136)\\
&\leq -\frac{1-s}{s} \liminf_{n\to\infty} \log \Bigg[\frac{1}{2} \mathrm{e}^{-\frac{s}{1-s} nR} \sum_{\mathbf{e}} P_E^n(\mathbf{e}) \Bigg(\sum_{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \epsilon \mathrm{e}^{-nR}} P_{A|E}^n(\mathbf{a}|\mathbf{e})^{1-s} \Bigg)^{\frac{1}{1-s}} \\
&\qquad + \frac{1}{2\epsilon^{\frac{s}{1-s}}} \sum_{\mathbf{e}} P_E^n(\mathbf{e}) \Bigg(\sum_{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) < \epsilon \mathrm{e}^{-nR}} P_{A|E}^n(\mathbf{a}|\mathbf{e})^{1-s} \Bigg)^{\frac{1}{1-s}} \Bigg]. \qquad (137)
\end{aligned}$$

In the following two steps, we evaluate the first and second terms in the $\liminf$ in (137).

In Step 2, we evaluate the second term in the $\liminf$ in (137) as it is simpler and provides the intuition and techniques for evaluating the first term. For this, we need to employ the Gärtner-Ellis theorem [48] (instead of Cramer's theorem). Doing so to the sequence of random variables $-\log P_{A|E}^n(A^n|\mathbf{e}) = \sum_{i=1}^n -\log P_{A|E}(A_i|e_i)$ with $\mathbf{e}$ of fixed type [14] and A^n with the memoryless distribution $P_{A|E}^n(\cdot|\mathbf{e})$, as will be shown in the following, we obtain

$$\begin{aligned}
&\lim_{n\to\infty} -\frac{1}{n} \log \sum_{\mathbf{e}} P_E^n(\mathbf{e}) \Bigg(\sum_{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) < \epsilon \mathrm{e}^{-nR}} P_{A|E}^n(\mathbf{a}|\mathbf{e}) \Bigg)^{\frac{1}{1-s}} \\
&\quad = \max_{t\geq 0} \frac{t}{1-s} \left(R - H_{1-t|1-s}(A|E|P_{AE}) \right), \qquad (138)
\end{aligned}$$

where $H_{1-t|1-s}(A|E|P_{AE})$ is the two-parameter conditional Rényi entropy defined in (22). To show (138), consider $\mathbf{e} \in \mathcal{T}_Q = \{\mathbf{e} \in \mathcal{E}^n : \text{type}(\mathbf{e}) = Q\}$. Then the Gärtner-Ellis theorem [48] yields that

$$P_{A^n|E^n=\mathbf{e}} \Big\{ \mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) < \epsilon \mathrm{e}^{-nR} \Big\} \doteq \exp \left(-n \max_{t\geq 0} \left[tR - \mathbb{E}_Q \log \sum_a P_{A|E}^{1-t}(a|E) \right] \right), \qquad (139)$$

where

$$\mathbb{E}_Q \log \sum_a P_{A|E}^{1-t}(a|E) = \sum_e Q(e) \log \sum_a P_{A|E}^{1-t}(a|e). \tag{140}$$

Let $\mathcal{P}_n(\mathcal{E})$ be the set of n-types with alphabet $\mathcal{E}$. Splitting the sum on the left-hand-side in (138) into the polynomially many n-types on $\mathcal{E}$, we obtain

$$\begin{aligned}
&\sum_{\mathbf{e}\in\mathcal{E}^n} P_E^n(\mathbf{e})\Big(\sum_{\mathbf{a}:P_{A|E}^n(\mathbf{a}|\mathbf{e})<\epsilon \mathrm{e}^{-nR}} P_{A|E}^n(\mathbf{a}|\mathbf{e})\Big)^{\frac{1}{1-s}} \\
&\doteq \sum_{Q\in\mathcal{P}_n(\mathcal{E})} P_E^n(\mathcal{T}_Q)\exp\left(-\frac{n}{1-s}\max_{t\ge 0}\left[tR-\mathbb{E}_Q\log\sum_a P_{A|E}^{1-t}(a|E)\right]\right) && (141)\\
&\doteq \max_{Q\in\mathcal{P}_n(\mathcal{E})} \exp\big(-nD(Q\|P_E)\big)\exp\left(-\frac{n}{1-s}\max_{t\ge 0}\left[tR-\mathbb{E}_Q\log\sum_a P_{A|E}^{1-t}(a|E)\right]\right) && (142)\\
&\doteq \exp\left(-n\min_Q\max_{t\ge 0}\left[\frac{tR}{1-s}-\frac{\mathbb{E}_Q\log\sum_a P_{A|E}^{1-t}(a|E)}{1-s}+D(Q\|P_E)\right]\right) && (143)\\
&= \exp\left(-n\max_{t\ge 0}\min_Q\left[\frac{tR}{1-s}-\frac{\mathbb{E}_Q\log\sum_a P_{A|E}^{1-t}(a|E)}{1-s}+D(Q\|P_E)\right]\right) && (144)
\end{aligned}$$

where (142) follows from the fact that $P_E^n(\mathcal{T}_Q)\doteq\exp(-nD(Q\|P_E))$ [14, Ch. 2], the swapping of $\min$ and $\max$ in (144) follows from the fact that the objective function is convex and concave in Q and t respectively, Q resides in a compact, convex set (the probability simplex) and t resides in a convex set $[0,\infty)$ (Sion's minimax theorem [60]). Now by straightforward calculus, the optimizing distribution for fixed t is

$$Q^*(e) = \frac{P_E(e)\big(\sum_a P_{A|E}^{1-t}(a|e)\big)^{\frac{1}{1-s}}}{Z_{s,t}} \tag{145}$$

where the normalizing constant (partition function)

$$Z_{s,t} := \sum_e P_E(e)\Big(\sum_a P_{A|E}^{1-t}(a|e)\Big)^{\frac{1}{1-s}}. \tag{146}$$

Plugging this into (144) we obtain

$$\sum_{\mathbf{e}\in\mathcal{E}^n} P_E^n(\mathbf{e})\Big(\sum_{\mathbf{a}:P_{A|E}^n(\mathbf{a}|\mathbf{e})<\epsilon \mathrm{e}^{-nR}} P_{A|E}^n(\mathbf{a}|\mathbf{e})\Big)^{\frac{1}{1-s}} \doteq \exp\left(-n\max_{t\ge 0}\frac{t}{1-s}\left[R-\frac{1-s}{t}Z_{s,t}\right]\right) \tag{147}$$

which then yields (138). Note that we have to use the Gärtner-Ellis theorem (and not Cramer's theorem) because the collection of random variables $\{-\log P_{A|E}(A_i|e_i): i=1,\dots,n\}$ is independent but *not identically distributed*.

In Step 3, we evaluate the first term in the $\liminf$ in (137) again by applying the Gärtner-Ellis theorem [48] to the sequence of random variables $\log P_{A|E}^n(A^n|\mathbf{e})=\sum_{i=1}^n \log P_{A|E}(A_i|e_i)$ with non-negative measure $P_{A|E}^n(\cdot|\mathbf{e})^{1-s}$, we have

$$\sum_{\mathbf{a}:P_{A|E}^n(\mathbf{a}|\mathbf{e})\ge\epsilon \mathrm{e}^{-nR}} P_{A|E}^n(\mathbf{a}|\mathbf{e})^{1-s} \doteq \exp\left(-n\max_{t\ge 0}\left[-tR-\mathbb{E}_Q\log\sum_a P_{A|E}^{1-(s-t)}(a|E)\right]\right) \tag{148}$$

where Q is the type of $\mathbf{e}$ and $\mathbb{E}_Q\log\sum_a P_{A|E}^{1-(s-t)}(a|E)$ is defined in (140). So by using a type partitioning argument of sequences $\mathbf{e}$ similarly to (141)–(147), we obtain

$$\begin{aligned}
&\sum_{\mathbf{e}} P_E^n(\mathbf{e})\Big(\sum_{\mathbf{a}:P_{A|E}^n(\mathbf{a}|\mathbf{e})\ge\epsilon \mathrm{e}^{-nR}} P_{A|E}^n(\mathbf{a}|\mathbf{e})^{1-s}\Big)^{\frac{1}{1-s}} \\
&\qquad\doteq \exp\left(-n\max_{t\ge 0}\left[-\frac{tR}{1-s}+\log\sum_e P_E(e)\Big(\sum_a P_{A|E}^{1-(s-t)}(a|e)\Big)^{\frac{1}{1-s}}\right]\right). && (149)
\end{aligned}$$

Consequently, considering the two different cases similarly to (132), we obtain

$$\lim_{n\to\infty} -\frac{1}{n}\log\left[\mathrm{e}^{-\frac{s}{1-s}nR}\sum_{\mathbf{e}} P_E^n(\mathbf{e})\left(\sum_{\mathbf{a}:P_{A|E}^n(\mathbf{a}|\mathbf{e})\geq\epsilon\mathrm{e}^{-nR}} P_{A|E}^n(\mathbf{a}|\mathbf{e})^{1-s}\right)^{\frac{1}{1-s}}\right]$$
$$= \begin{cases} \frac{s}{1-s}(R - H_{1-s}^{\uparrow}(A|E|P_{AE})) & \text{if } R \geq \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t|1-s}(A|E|P_{AE})\big|_{t=s} \\ \max_{t\in[0,s]} \frac{t}{1-s}(R - H_{1-t|1-s}(A|E|P_{AE})) & \text{if } R \leq \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t|1-s}(A|E|P_{AE})\big|_{t=s} \end{cases}. \tag{150}$$

In Step 4, we put together the asymptotic evaluations in (138) and (150) in the bound in (137). We observe that (138) is not smaller than (150). Thus, the former dominates the exponential behavior, and plugging (138) into (137), we obtain

$$\limsup_{n\to\infty} \frac{1}{n} C_{1-s}^{\uparrow}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})$$
$$\leq \begin{cases} R - H_{1-s}^{\uparrow}(A|E|P_{AE}) & \text{if } R \geq \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t|1-s}(A|E|P_{AE})\big|_{t=s} \\ \max_{t\in[0,s]} \frac{t}{s}(R - H_{1-t|1-s}(A|E|P_{AE})) & \text{if } R \leq \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t|1-s}(A|E|P_{AE})\big|_{t=s} \end{cases}. \tag{151}$$

Finally in Step 5, we show that the transition rate in (151)

$$\frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t|1-s}(A|E|P_{AE})\big|_{t=s} = \hat{R}_{-s}^{\uparrow} \tag{152}$$

as follows. Since $\max_t H_{1-s|1-t}(A|E|P_{AE}) = H_{1-s|1-s}(A|E|P_{AE})$ (i.e., the maximum is attained at $t = s$),

$$\frac{\mathrm{d}}{\mathrm{d}t} sH_{1-s|1-t}(A|E|P_{AE})\big|_{t=s} = s\frac{\mathrm{d}}{\mathrm{d}t} H_{1-s|1-t}(A|E|P_{AE})\big|_{t=s} = 0. \tag{153}$$

Hence, choosing $t_1 = t_2 = t$, we have

$$\hat{R}_{-s}^{\uparrow} = \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t}^{\uparrow}(A|E|P_{AE})\big|_{t=s} \tag{154}$$
$$= \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t|1-t}(A|E|P_{AE})\big|_{t=s} \tag{155}$$
$$= \frac{\mathrm{d}t_1}{\mathrm{d}t}\frac{\partial}{\partial t_1} t_1 H_{1-t_1|1-t_2}(A|E|P_{AE})\big|_{t=s} + \frac{\mathrm{d}t_2}{\mathrm{d}t}\frac{\partial}{\partial t_2} t_1 H_{1-t_1|1-t_2}(A|E|P_{AE})\big|_{t=s} \tag{156}$$
$$= \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t|1-s}(A|E|P_{AE})\big|_{t=s} + \frac{\mathrm{d}}{\mathrm{d}t} sH_{1-s|1-t}(A|E|P_{AE})\big|_{t=s} \tag{157}$$
$$= \frac{\mathrm{d}}{\mathrm{d}t} tH_{1-t|1-s}(A|E|P_{AE})\big|_{t=s}, \tag{158}$$

where (155) follows from the relation in (52), (156) follows from the chain rule, and (158) follows from (153). The relations in (151) and (152) complete the justification of the upper bound of (50).

2) Converse Parts: For the converse, we do not consider the common randomness X_n (i.e., $X_n = \emptyset$) since the bound must hold for *all* (not just ϵ-almost universal$_2$) hash functions f_{X_n}. This statement applies to the proofs of all converse bounds in the sequel.

a) Proofs of the lower bounds of (47) *and* (48)*:* The lower bounds to (47) and (48) can be easily obtained by using the data processing inequalities for Rényi conditional entropies and their Gallager-type counterparts in (26)–(28).

b) Proof of the lower bound of (49)*:* Now for (49), we note that when $R \geq \hat{R}_{-s}$, we have

$$s(R - H_{1-s}(A|E|P_{AE})) \leq \max_{t\in[0,s]} t(R - H_{1-t}(A|E|P_{AE})) \tag{159}$$

and when $R \leq \hat{R}_{-s}$, equality holds since $t = s$ attains the maximum. Fix $t \in [0, s]$. From (109), we obtain the bound

$$\mathrm{e}^{-sC_{1-s}(f(A^n)|E^n|P_{AE}^n)} \leq c^{-s}\sum_{\mathbf{e}} P_{E^n}(\mathbf{e}) \sum_{\mathbf{a}:P_{A|E}^n(\mathbf{a}|\mathbf{e})\geq\frac{c}{M_n}} P_{A|E}^n(\mathbf{a}|\mathbf{e})^{1-s} M_n^{-s}$$
$$+ 2^{\frac{s}{1-s}} s^{\frac{s}{1-s}}(1-s) P_{AE}^n\left\{(\mathbf{a},\mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) < \frac{c}{M_n}\right\} \tag{160}$$
$$\leq c^{-s}\mathrm{e}^{sH_{1-s}(A^n|E^n|P_{AE}^n)} M_n^{-s} + 2^{\frac{s}{1-s}} s^{\frac{1}{1-s}-1}(1-s)\mathrm{e}^{tH_{1-t}(A^n|E^n|P_{AE}^n)}\left(\frac{M_n}{c}\right)^{-t}, \tag{161}$$

where in (161), we upper bounded the probability in the second term using Markov's inequality, i.e., for any $t \in [0, s]$,

$$P^n_{AE}\Big\{(\mathbf{a},\mathbf{e}) : P^n_{A|E}(\mathbf{a}|\mathbf{e}) < \frac{c}{M_n}\Big\} = P^n_{AE}\Big\{(\mathbf{a},\mathbf{e}) : \mathrm{e}^{-t\log P^n_{A|E}(\mathbf{a}|\mathbf{e})} > \mathrm{e}^{-t\log\frac{c}{M_n}}\Big\} \tag{162}$$

$$\leq \frac{\mathbb{E}_{P^n_{AE}}\left[\mathrm{e}^{-t\log P^n_{A|E}(A^n|E^n)}\right]}{\mathrm{e}^{-t\log\frac{c}{M_n}}} \tag{163}$$

$$= \mathrm{e}^{tH_{1-t}(A^n|E^n|P^n_{AE})}\Big(\frac{M_n}{c}\Big)^{-t}. \tag{164}$$

Put $c = 1$ in (161). We then obtain the lower bound to (49) by applying (159) and its equality version for $R \leq \hat{R}_{-s}$.

c) Proof of the lower bound of (50)*:* Finally, (50) can be obtained by evaluating (110) as follows:

$$\begin{aligned}
&\mathrm{e}^{-\frac{s}{1-s}C^{\uparrow}_{1-s}(f(A^n)|E^n|P^n_{AE})}\\
&\leq 2^{\frac{s}{1-s}}\sum_{\mathbf{e}}\Bigg[P_{E^n}(\mathbf{e})\Big(c^{-s}\sum_{\mathbf{a}:P^n_{A|E}(\mathbf{a}|\mathbf{e})\geq\frac{c}{M_n}}P^n_{A|E}(\mathbf{a}|\mathbf{e})^{1-s}M_n^{-s}\Big)^{\frac{1}{1-s}}\\
&\qquad+\Big(2^{\frac{s}{1-s}}s^{\frac{s}{1-s}}(1-s)P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P^n_{A|E}(\mathbf{a}|\mathbf{e}) < \frac{c}{M_n}\Big\}\Big)^{\frac{1}{1-s}}\Bigg] &(165)\\
&\leq 2^{\frac{s}{1-s}}\Bigg[c^{-s}\mathrm{e}^{\frac{s}{1-s}H^{\uparrow}_{1-s}(A^n|E^n|P^n_{AE})}M_n^{-\frac{s}{1-s}}\\
&\qquad+2^{\frac{s}{(1-s)^2}}s^{\frac{s}{(1-s)^2}}(1-s)^{\frac{1}{1-s}}\sum_{\mathbf{e}}P^n_E(\mathbf{e})\mathrm{e}^{\frac{t}{1-s}H_{1-t}(A^n|P_{A^n|E^n=\mathbf{e}})}\Big(\frac{M_n}{c}\Big)^{-\frac{t}{1-s}}\Bigg] &(166)\\
&= 2^{\frac{s}{1-s}}\Bigg[c^{-s}\mathrm{e}^{\frac{s}{1-s}H^{\uparrow}_{1-s}(A^n|E^n|P^n_{AE})}M_n^{-\frac{s}{1-s}}+2^{\frac{s}{(1-s)^2}}s^{\frac{s}{(1-s)^2}}(1-s)^{\frac{1}{1-s}}\mathrm{e}^{\frac{t}{1-s}H_{1-t|1-s}(A^n|E^n|P^n_{AE})}\Big(\frac{M_n}{c}\Big)^{-\frac{t}{1-s}}\Bigg] &(167)
\end{aligned}$$

with $s \geq t \geq 0$. For brevity, let $\beta_s := 2^{\frac{s}{(1-s)^2}}s^{\frac{s}{(1-s)^2}}(1-s)^{\frac{1}{1-s}}$ be a function that only depends on s. By taking the logarithm of (167), normalizing by n, and using $M_n = \mathrm{e}^{nR}$, we obtain

$$\begin{aligned}
&\frac{1}{n}C^{\uparrow}_{1-s}(f(A^n)|E^n|P^n_{AE})\\
&\geq -\frac{1-s}{sn}\log\Bigg[2^{\frac{s}{1-s}}\Bigg[c^{-s}\mathrm{e}^{\frac{s}{1-s}H^{\uparrow}_{1-s}(A^n|E^n|P^n_{AE})}M_n^{-\frac{s}{1-s}}+\beta_s\mathrm{e}^{\frac{t}{1-s}H_{1-t|1-s}(A^n|E^n|P^n_{AE})}\Big(\frac{M_n}{c}\Big)^{-\frac{t}{1-s}}\Bigg]\Bigg] &(168)\\
&= -\frac{1-s}{sn}\log\Bigg[2^{\frac{s}{1-s}}\Bigg[c^{-s}\mathrm{e}^{-n\frac{s}{1-s}(R-H^{\uparrow}_{1-s}(A|E|P_{AE}))}+\beta_s c^{\frac{t}{1-s}}\mathrm{e}^{-n\frac{t}{1-s}(R-H_{1-t|1-s}(A|E|P_{AE}))}\Bigg]\Bigg]. &(169)
\end{aligned}$$

Now similarly to (159), we have

$$s(R - H^{\uparrow}_{1-s}(A|E|P_{AE})) \leq \max_{t\in[0,s]} t(R - H_{1-t|1-s}(A|E|P_{AE})). \tag{170}$$

Thus when $R \geq \hat{R}^{\uparrow}_{-s}$, the first term in (169) dominates (exponent is not larger), resulting in the first clause in (50), namely $R - H^{\uparrow}_{1-s}(A|E|P_{AE})$. On the other hand, when $R \leq \hat{R}^{\uparrow}_{-s}$, the second term in (169) dominates. To complete the argument, we optimize over $t \in [0, s]$ to obtain the second clause in (50), namely $\max_{t\in[0,s]}\frac{t}{s}(R - H_{1-t|1-s}(A|E|P_{AE}))$. This completes the proof of the lower bound of (50).

B. *Proof of Theorem 2*

1) *Direct Parts:*

a) Proof of the lower bound of (55)*:* We note, per the discussion following Theorem 2, that

$$C_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) = O(n), \quad \text{and} \tag{171}$$

$$C^{\uparrow}_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) = O(n). \tag{172}$$

Thus, the exponents are lower bounded by zero, explaining the $|\cdot|^+$ in (55) and (57).

Now for the non-trivial (non-zero) lower bound on the exponents, we employ (102) with $\epsilon = 1$ and $t \in [s, 1]$. We recall that $M_n = \mathrm{e}^{nR}$. Now we have

$$\begin{aligned} & C_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \\ & \leq C_{1+t}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \quad (173) \\ & \leq \frac{1}{t} \log \left(1 + M_n^t \mathrm{e}^{-tH_{1+t}(A^n|E^n|P_{AE}^n)}\right) \quad (174) \\ & \leq \frac{1}{t} M_n^t \mathrm{e}^{-ntH_{1+t}(A|E|P_{AE})}. \quad (175) \end{aligned}$$

Taking the logarithm and optimizing over $t \in [s, 1]$, we obtain the lower bound to (55).

b) Proof of the lower bound of (57)*:* Similarly, applying (106) to the case $t \in [s, 1]$, we obtain

$$\begin{aligned} & C^{\uparrow}_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \\ & \leq C^{\uparrow}_{1+t}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \quad (176) \\ & \leq \frac{1+t}{t} \log \left(1 + \frac{1}{1+t} M_n^t \mathrm{e}^{-tH_{1+t}(A^n|E^n|P_{AE}^n)}\right) \quad (177) \\ & \leq \frac{1}{t} M_n^t \mathrm{e}^{-ntH_{1+t}(A|E|P_{AE})} \quad (178) \end{aligned}$$

which implies the lower bound to (57) upon optimizing over $t \in [s, 1]$.

c) Proofs of the lower bounds of (56) *and* (58)*:* For the $-s$ versions in (56) and (58), we simply note that

$$C_{1+s} \geq C_{1-s'} \tag{179}$$

$$C^{\uparrow}_{1+s} \geq C^{\uparrow}_{1-s'} \tag{180}$$

for any $s, s' \in [0, 1]$ because as mentioned in Section II-A (after (15) and (20) respectively), H_{1+s} and $H^{\dagger}_{1+s}$ are monotonically decreasing in s. Thus, we have

$$\liminf_{n\to\infty} -\frac{1}{n} \log C_{1-s'}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \geq \liminf_{n\to\infty} -\frac{1}{n} \log C_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \tag{181}$$

$$\liminf_{n\to\infty} -\frac{1}{n} \log C^{\uparrow}_{1-s'}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})) \geq \liminf_{n\to\infty} -\frac{1}{n} \log C^{\uparrow}_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})). \tag{182}$$

Combining these statements with the bounds derived in (173)–(178) completes the proof of (56) and (58).

2) Converse Parts:

a) Proof of the upper bound of (55)*:* For the converse, we first show the upper bound to (55). Choose a constant c_0 satisfying $c_0^s > 1 + s$. Recall the definition of $\hat{R}_s$ in (24). Now assume that $R \geq \hat{R}_s$. We claim that

$$\lim_{n\to\infty} -\frac{1}{n} \log \left[P_{AE}^n \left\{(\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-n\hat{R}_s}\right\} c_0^s \mathrm{e}^{-sn\hat{R}_s} \mathrm{e}^{snR}\right] = sH_{1+s}(A|E|P_{AE}) - sR. \tag{183}$$

This is justified as follows. We know from Cramér's theorem [48] that

$$P_{AE}^n \left\{(\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-n\hat{R}_s}\right\} \doteq \exp\left(-n \max_{t\geq 0} \left\{tH_{1+t}(A|E|P_{AE}) - t\hat{R}_s\right\}\right). \tag{184}$$

The maximum appeared in right-hand-side of (184) is attained when the derivative of $tH_{1+t}(A|E|P_{AE}) - t\hat{R}_s$ is zero because $tH_{1+t}(A|E|P_{AE})$ is concave in t. Hence, the real number t satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} tH_{1+t}(A|E|P_{AE}) = \hat{R}_s, \tag{185}$$

which implies $t = s \geq 0$ due to the definition of $\hat{R}_s$ in (24) and the strictly decreasing nature of $\hat{R}_s$. As a result,

$$P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-n\hat{R}_s} \right\} \doteq \exp\left(-n \left\{ sH_{1+s}(A|E|P_{AE}) - s\hat{R}_s \right\} \right). \tag{186}$$

Plugging this into the left-hand-side of (183) yields the claim. The one-shot bound in (115) with $c = c_0 \mathrm{e}^{-n\hat{R}_s + nR} > 1$ implies that

$$\begin{aligned}
&\mathrm{e}^{sC_{1+s}(f(A^n)|E^n|P_{AE}^n)} \\
&\geq P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-n\hat{R}_s} \right\} c_0^s \mathrm{e}^{-n\hat{R}_s} \mathrm{e}^{snR} \\
&\qquad + 1 - (1+s) P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-n\hat{R}_s} \right\} &(187)\\
&= 1 + \left(c_0^s \mathrm{e}^{-sn\hat{R}_s} \mathrm{e}^{snR} - 1 - s \right) P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-n\hat{R}_s} \right\}. &(188)
\end{aligned}$$

Hence, taking the logarithm of (188), employing the lower bound $\log(1+b) \geq b - \frac{b^2}{2}$, the large-deviations result (184), and the fact that $\lim_{n \to \infty} \frac{1}{n} \log(c_0^s \mathrm{e}^{-sn\hat{R}_s + snR} - 1 - s) = s(-\hat{R}_s + R)$, we obtain

$$sC_{1+s}(f(A^n)|E^n|P_{AE}^n) \dot{\geq} \exp\left[-ns(H_{1+s}(A|E|P_{AE}) - R) \right]. \tag{189}$$

Finally, we obtain the upper bound to (55) by taking another logarithm and normalizing by n.

For the other case $R \leq \hat{R}_s$, we claim that

$$\max_{t \geq s} \{ tH_{1+t}(A|E|P_{AE}) - tR \} = \max_{t \geq 0} \{ tH_{1+t}(A|E|P_{AE}) - tR \}. \tag{190}$$

This is because by the strict concavity of $t \mapsto tH_{1+t}$, the map $s \mapsto \hat{R}_s$ is strictly decreasing. So for $R \leq \hat{R}_s$ the maximum on the right-hand-side of (190) is attained at some $t \geq s$. This is also reflected in Fig. 5. Thus, (112) implies that

$$\begin{aligned}
&\mathrm{e}^{sC_{1+s}(f(A^n)|E^n|P_{AE}^n)} \\
&\geq P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-nR} \right\} c_0^s \\
&\qquad + \left(1 - P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-nsR} \right\} \right)^{1+s} &(191)\\
&\geq 1 + (c_0^s - 1 - s) P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-nsR} \right\}. &(192)
\end{aligned}$$

Hence,

$$\begin{aligned}
&C_{1+s}(f(A^n)|E^n|P_{AE}^n) \\
&= \frac{1}{s} \log(\mathrm{e}^{sC_{1+s}(f(A^n)|E^n|P_{AE}^n)}) &(193)\\
&\geq \frac{1}{s} \log\left(1 + (c_0^s - 1 - s) P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-nsR} \right\} \right) &(194)\\
&\doteq \frac{c_0^s - 1 - s}{s} P_{AE}^n \left\{ (\mathbf{a}, \mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0 \mathrm{e}^{-nsR} \right\}. &(195)
\end{aligned}$$

By combining the asymptotic evaluation using Cramér's theorem in (184) and the equality in (190), we see that for $R \leq \hat{R}_s$, we also obtain the upper bound to (55). This completes the proof.

b) Proof of the upper bound of (57)*:* The proof of the upper bound to (57) is similar and we present the details here. Similarly to the above proof, choose $c = c_0 \mathrm{e}^{-n\hat{R}_s^{\uparrow} + nR} > 1$ and the constant $c_0^s > 1 + s$. Assume that

$R \geq \hat{R}_s^{\uparrow}$, where $\hat{R}_s^{\uparrow}$ is defined in (25). Then the one-shot bound in (116) implies that

$$\begin{aligned}
&\mathrm{e}^{\frac{s}{1+s}C_{1+s}^{\uparrow}(f(A^n)|E^n|P_{AE}^n)} \\
&\geq \sum_{\mathbf{e}} P_E^n(\mathbf{e})\Big(P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\} c_0^s\mathrm{e}^{-sn\hat{R}_s^{\uparrow}}\mathrm{e}^{snR} \\
&\qquad + \Big(1 - P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\}\Big)^{1+s}\Big)^{\frac{1}{1+s}} && (196)\\
&\geq \sum_{\mathbf{e}} P_E^n(\mathbf{e})\Big(P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\} c_0^s\mathrm{e}^{-sn\hat{R}_s^{\uparrow}}\mathrm{e}^{snR} \\
&\qquad + 1 - (1+s)P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\}\Big)^{\frac{1}{1+s}} && (197)\\
&= \sum_{\mathbf{e}} P_E^n(\mathbf{e})\left(1 + \big(c_0^s\mathrm{e}^{-sn\hat{R}_s^{\uparrow}}\mathrm{e}^{snR} - (1+s)\big)P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\}\right)^{\frac{1}{1+s}} && (198)\\
&\doteq \sum_{\mathbf{e}} P_E^n(\mathbf{e})\Big(1 + \frac{c_0^s\mathrm{e}^{-sn\hat{R}_s^{\uparrow}}\mathrm{e}^{snR} - (1+s)}{1+s}P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\}\Big) && (199)\\
&= 1 + \frac{c_0^s\mathrm{e}^{-sn\hat{R}_s^{\uparrow}}\mathrm{e}^{snR} - (1+s)}{1+s}\sum_{\mathbf{e}} P_E^n(\mathbf{e})P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\}, && (200)
\end{aligned}$$

where (199) follows from the fact that $\log[(1+a)^t] = t[a + O(a^2)]$ for $a \downarrow 0$. Hence,

$$\begin{aligned}
&C_{1+s}^{\uparrow}(f(A^n)|E^n|P_{AE}^n) \\
&= \frac{1+s}{s}\log\left(\mathrm{e}^{\frac{s}{1+s}C_{1+s}^{\uparrow}(f(A^n)|E^nX_n|P_{AE}^n\times P_{X_n})}\right) && (201)\\
&\geq \frac{1+s}{s}\log\left(1 + \frac{c_0^s\mathrm{e}^{-sn\hat{R}_s^{\uparrow}}\mathrm{e}^{snR} - (1+s)}{1+s}\sum_{\mathbf{e}} P_E^n(\mathbf{e})P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\}\right) && (202)\\
&\doteq \frac{c_0^s\mathrm{e}^{-sn\hat{R}_s^{\uparrow}}\mathrm{e}^{snR} - (1+s)}{s}\sum_{\mathbf{e}} P_E^n(\mathbf{e})P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}: P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\}, && (203)\\
&= \frac{c_0^s\mathrm{e}^{-sn\hat{R}_s^{\uparrow}}\mathrm{e}^{snR} - (1+s)}{s}P_{AE}^n\left\{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq c_0\mathrm{e}^{-n\hat{R}_s^{\uparrow}}\right\}, && (204)
\end{aligned}$$

where (204) follows from $\log(1+a) = a + O(a^2)$ as $a \downarrow 0$ and the fact that the summation in (202) vanishes as n grows. Combining (183) and (204) yields the upper bound to (57) for $R \geq \hat{R}_s^{\uparrow}$. A similar calculation for the case $R \leq \hat{R}_s^{\uparrow}$ also yields the the same upper bound to (57).

c) Proof of the upper bound of (56)*:* We choose the constant c such that $(1-s) > c^{-s}$. We apply Cramer's Theorem [48] to the sequence of random variables $\log P_{A|E}^n(A^n|E^n)$. Then,

$$\lim_{n\to\infty} -\frac{1}{n}\log P_{AE}^n\left\{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{\mathrm{e}^{nR}}\right\} = \max_{t\geq 0}\left\{tH_{1+t}(A|E|P_{AE}) - tR\right\}. \tag{205}$$

The one-shot bound in (112) implies that

$$\begin{aligned}
&\mathrm{e}^{-sC_{1-s}(f(A^n)|E^n|P_{AE}^n)} \\
&\leq P_{AE}^n\left\{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\right\}c^{-s} + P_{AE}^n\left\{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) < \frac{c}{M}\right\}^{1-s} && (206)\\
&= P_{AE}^n\left\{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\right\}c^{-s} + \left(1 - P_{AE}^n\left\{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\right\}\right)^{1-s} && (207)\\
&\leq 1 - ((1-s) - c^{-s})P_{AE}^n\left\{(\mathbf{a},\mathbf{e}): P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\right\}. && (208)
\end{aligned}$$

Thus,

$$\begin{aligned}
&C_{1-s}(f(A^n)|E^n|P_{AE}^n) \\
&= -\frac{1}{s}\log\left[\mathrm{e}^{-sC_{1-s}(f(A^n)|E^n|P_{AE}^n)}\right] && (209)\\
&\geq -\frac{1}{s}\log\left[1-((1-s)-c^{-s})P_{AE}^n\Big\{(\mathbf{a},\mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}\right] && (210)\\
&\geq \frac{1}{s}((1-s)-c^{-s})P_{AE}^n\Big\{(\mathbf{a},\mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}, && (211)
\end{aligned}$$

where the final step uses the inequality $\log(1-t) \leq -t$. Combining the limiting statement in (205) and the bound in (211), we have the upper bound to (56).

d) Proof of the upper bound of (58)*:* When $\sum_{\mathbf{e}} P_E^n(\mathbf{e})P_{A^n|E^n=\mathbf{e}}\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\}$ is exponentially small, (113) implies that

$$\begin{aligned}
&\mathrm{e}^{-\frac{s}{1-s}C_{1-s}^{\uparrow}(f(A^n)|E^n|P_{AE}^n)} \\
&\leq \sum_{\mathbf{e}} P_E^n(\mathbf{e})\bigg(P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}c^{-s} \\
&\qquad + P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) < \frac{c}{M}\Big\}^{1-s}\bigg)^{\frac{1}{1-s}} && (212)\\
&= \sum_{\mathbf{e}} P_E^n(\mathbf{e})\bigg(c^{-s}P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\} \\
&\qquad + \Big(1-P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}\Big)^{1-s}\bigg)^{\frac{1}{1-s}} && (213)\\
&\leq \sum_{\mathbf{e}} P_E^n(\mathbf{e})\left(1-(1-s-c^{-s})P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}\right)^{\frac{1}{1-s}} && (214)\\
&\doteq \sum_{\mathbf{e}} P_E^n(\mathbf{e})\left(1-\frac{1-s-c^{-s}}{1-s}P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}\right) && (215)\\
&= 1-\frac{1-s-c^{-s}}{1-s}\sum_{\mathbf{e}} P_E^n(\mathbf{e})P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}, && (216)
\end{aligned}$$

where (215) follows from the same reasoning as (199). Thus,

$$\begin{aligned}
&C_{1-s}^{\uparrow}(f(A^n)|E^n|P_{AE}^n) = -\frac{1-s}{s}\log\left[\mathrm{e}^{-\frac{s}{1-s}C_{1-s}^{\uparrow}(f(A^n)|E^n|P_{AE}^n)}\right] && (217)\\
&\geq -\frac{1-s}{s}\log\left[1-\frac{1-s-c^{-s}}{1-s}\sum_{\mathbf{e}} P_E^n(\mathbf{e})P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}\right] && (218)\\
&\geq \frac{1-s}{s}\cdot\frac{1-s-c^{-s}}{1-s}\sum_{\mathbf{e}} P_E^n(\mathbf{e})P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\} && (219)\\
&= \frac{1-s-c^{-s}}{s}\sum_{\mathbf{e}} P_E^n(\mathbf{e})P_{A^n|E^n=\mathbf{e}}\Big\{\mathbf{a} : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M}\Big\}. && (220)
\end{aligned}$$

Combining (205) and (220), we have the upper bound to (58).

C. Proof of Theorem 3

1) Direct Parts:

a) Proof of upper bounds for Case (A): First, we prove the upper bounds for Case (A) where the Rényi parameter $\alpha = 1 + s$ for $s \in (0, 1]$. Substituting $\mathrm{e}^{nH_{1+s}(A|E|P_{AE})+\sqrt{n}L}$ into M_n in the chain of inequalities in (118)–(119), we obtain, for the class of ϵ-almost universal$_2$ hash functions f_{X_n}, that

$$C_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \leq \frac{1}{s}\log\left(\epsilon^s + \mathrm{e}^{s\sqrt{n}L}\right). \tag{221}$$

Set ϵ to be a constant (not varying with n). Normalizing by $\sqrt{n}$ and taking the $\limsup$ as $n \to \infty$ yields the upper bound to (68).

In an exactly analogous way, the upper bound to (69) can be shown by substituting $\mathrm{e}^{nH^{\uparrow}_{1+s}(A|E|P_{AE})+\sqrt{n}L}$ into M_n in the chain of inequalities in (176)–(178).

Substituting $\mathrm{e}^{nH_{1+s}(A|E|P_{AE})+\sqrt{n}L}$ into M_n in the chain of inequalities in (173)–(175) with $t = s$, we obtain

$$C_{1+s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \leq M_n^s \mathrm{e}^{-snH_{1+s}(A|E|P_{AE})} = \mathrm{e}^{s\sqrt{n}L} \tag{222}$$

which implies the upper bound to (70) after we take the logarithm, normalize both sides by $\sqrt{n}$ and take the $\limsup$ as $n \to \infty$.

In an exactly analogous way, the upper bound to (71) can be shown by substituting $\mathrm{e}^{nH^{\uparrow}_{1+s}(A|E|P_{AE})+\sqrt{n}L}$ into M_n in the chain of inequalities in (120)–(121). This completes the proof for the direct part of Case (A) of Theorem 3.

b) Proof of upper bound for Case (B): Case (B) follows from four distinct steps, detailed in each of the following paragraphs.

In Step 1, we fix any function $f : \mathcal{A}^n \to \{1, \ldots, \|f\|\}$. We partition the space $\mathcal{A}^n \times \mathcal{E}^n$ into pairs of sequences of the same joint type [14]. Let Q_{AE} denote a generic joint type on $\mathcal{A} \times \mathcal{E}$. Let $U^{(Q_{AE})}$ be the uniform distribution over the type class $\mathcal{T}_{Q_{AE}} \subset \mathcal{A}^n \times \mathcal{E}^n$. Let

$$U^{(Q_{AE})}_{f(A^n),E^n}(i, \mathbf{e}) := \sum_{\mathbf{a}: f(\mathbf{a})=i} U^{(Q_{AE})}(\mathbf{a}, \mathbf{e}) \tag{223}$$

be the distribution on $\{1, \ldots, \|f\|\} \times \mathcal{E}^n$ when the hash function f is applied to the variable A^n and denote

$$U^{(Q_{AE})}_{E^n}(\mathbf{e}) := \sum_{i=1}^{\|f\|} U^{(Q_{AE})}_{f(A^n),E^n}(i, \mathbf{e}) \tag{224}$$

as its $\mathcal{E}^n$-marginal. Because the probability of pairs of sequences of the same joint type have the same P_{AE}^n-probability, we can write

$$P_{f(A^n),E^n}(i, \mathbf{e}) = \sum_{Q_{AE} \in \mathcal{P}_n(\mathcal{A}\times\mathcal{E})} P_{AE}^n(\mathcal{T}_{Q_{AE}}) U^{(Q_{AE})}_{f(A^n),E^n}(i, \mathbf{e}). \tag{225}$$

By using (225), we have

$$\begin{aligned}
C_1(f(A^n)|E^n|P_{AE}^n) &= D(P_{f(A^n),E^n} \| P_{\mathrm{mix},f(\mathcal{A}^n)} \times P_{E^n}) && (226)\\
&\leq \sum_{Q_{AE} \in \mathcal{P}_n(\mathcal{A}\times\mathcal{E})} P_{AE}^n(\mathcal{T}_{Q_{AE}}) D\left(U^{(Q_{AE})}_{f(A^n),E^n} \middle\| P_{\mathrm{mix},f(\mathcal{A}^n)} \times U^{(Q_{AE})}_{E^n}\right) && (227)\\
&= \sum_{Q_{AE} \in \mathcal{P}_n(\mathcal{A}\times\mathcal{E})} P_{AE}^n(\mathcal{T}_{Q_{AE}}) C_1\left(f(A^n)|E^n \middle| U^{(Q_{AE})}_{A^n E^n}\right) && (228)\\
&\leq \sum_{Q_{AE} \in \mathcal{P}_n(\mathcal{A}\times\mathcal{E})} P_{AE}^n(\mathcal{T}_{Q_{AE}}) C_2\left(f(A^n)|E^n \middle| U^{(Q_{AE})}_{A^n E^n}\right). && (229)
\end{aligned}$$

where (227) follows from the fact that relative entropy is convex, (228) follows from the definition of C_1, and (229) follows from the fact that $s \mapsto C_{1+s}$ is monotonically non-decreasing.

In Step 2, we regard f as a universal$_2$ hash function f_{X_n}. Thus, (229) implies that

$$\begin{aligned}
C_1(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) &= \mathbb{E}_{X_n}\left[C_1(f_{X_n}(A^n)|E^n|P_{AE}^n)\right] && (230)\\
&\leq \sum_{Q_{AE} \in \mathcal{P}_n(\mathcal{A}\times\mathcal{E})} P_{AE}^n(\mathcal{T}_{Q_{AE}}) \mathbb{E}_{X_n}\left[C_2\left(f_{X_n}(A^n)|E^n \middle| U^{(Q_{AE})}_{A^n E^n}\right)\right]. && (231)
\end{aligned}$$

Let $\mathcal{T}_{Q_{A|E}}(\mathbf{e}) := \{\mathbf{a} : (\mathbf{a}, \mathbf{e}) \in \mathcal{T}_{Q_{AE}}\}$ be the conditional type class of $Q_{A|E}$ given $\mathbf{e}$, also known as the $Q_{A|E}$-shell. By the method of types [14, Ch. 2], we know that for $\mathbf{e}$ of type Q_E,

$$\log\left|\mathcal{T}_{Q_{A|E}}(\mathbf{e})\right| = nH(A|E|Q_{AE}) + O(\log n). \tag{232}$$

By using the fact that $\mathrm{e}^{-H_2(A|E)}$ is the conditional collision probability (i.e., $\mathrm{e}^{-H_2(A|E)} = \sum_e P_E(e)P_{AA'|E=e}\{(a,a') : a = a'\}$ where A, A' are conditionally independent and identically distributed given E),

$$\begin{aligned}
\mathrm{e}^{-H_2(A|E|U^{(Q_{AE})})} &= \sum_{\mathbf{e}} U_{E^n}^{(Q_{AE})}(\mathbf{e}) \sum_{\mathbf{a}\in\mathcal{T}_{Q_{A|E}}(\mathbf{e})} \frac{1}{|\mathcal{T}_{Q_{A|E}}(\mathbf{e})|^2} && (233)\\
&= \sum_{\mathbf{e}} U_{E^n}^{(Q_{AE})}(\mathbf{e}) \frac{1}{|\mathcal{T}_{Q_{A|E}}(\mathbf{e})|} && (234)\\
&= \mathrm{e}^{-nH(A|E|Q_{AE})+O(\log n)}. && (235)
\end{aligned}$$

Furthermore, by a Taylor expansion of $H(A|E|Q_{AE})$ around P_{AE} as in the rate redundancy lemma [51], [61], we have

$$H(A|E|Q_{AE}) = H(A|E|P_{AE}) + \sum_{a,e}(Q_{AE}(a,e) - P_{AE}(a,e))h_{A|E}(a|e) + O\big(\|Q_{AE} - P_{AE}\|^2\big) \tag{236}$$

where the *conditional entropy density* $h_{A|E}(a|e)$ is defined as

$$h_{A|E}(a|e) := \log\frac{1}{P_{A|E}(a|e)} \tag{237}$$

and $\|Q - P\| = \sum_{z\in\mathcal{Z}} |Q(z) - P(z)|$ is the variational distance between Q and P. For brevity, we denote the $\sqrt{n}$-scaled version of the second term in (236) as

$$b_n(Q_{AE}) := \sqrt{n}\left(\sum_{a,e}(Q_{AE}(a,e) - P_{AE}(a,e))h_{A|E}(a,e)\right). \tag{238}$$

If Q_{AE} is a random type formed from n independent copies of P_{AE},

$$b_n(Q_{AE}) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} h_{A|E}(A_i|E_i) - H(A|E|P_{AE})\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, V(A|E|P_{AE})\right) \tag{239}$$

by the central limit theorem. That is, $b_n(Q_{AE})$ converges in distribution to the Gaussian $\mathcal{N}\left(0, V(A|E|P_{AE})\right)$.

In Step 3, we first fix $\delta > 0$. Applying the universal$_2$ property of the universal$_2$ hash function f_{X_n} to the collision relative entropy (see (118)–(119) with $\epsilon = s = 1$), and combining the above notations and bounds, we obtain for all $\mathbf{e} \in \mathcal{T}_{Q_E}$ and all n large enough (depending on δ) that

$$\begin{aligned}
&\mathbb{E}_{X_n}\left[C_2\left(f_{X_n}(A^n)|E^n \,\middle|\, U_{A^nE^n}^{(Q_{AE})}\right)\right]\\
&\quad= \mathbb{E}_{X_n}\left[\log M - H_2\left(f_{X_n}(A^n)|E^n \,\middle|\, U_{A^nE^n}^{(Q_{AE})}\right)\right] && (240)\\
&\quad\le \log\left(1 + M_n \mathrm{e}^{-nH(A|E|Q_{AE})+O(\log n)}\right) && (241)\\
&\quad\le \log\left(1 + \exp\left[\sqrt{n}(L - b_n(Q_{AE}) + o(b_n(Q_{AE})) + O(\log n)\right]\right) && (242)\\
&\quad\le \begin{cases} \sqrt{n}\big(L - b_n(Q_{AE}) + o(b_n(Q_{AE}))\big) + O(\log n) & b_n(Q_{AE}) \le L + \delta \\ \mathrm{e}^{-\delta\sqrt{n}/2} & b_n(Q_{AE}) > L + \delta \end{cases}, && (243)
\end{aligned}$$

where (240) follows from the definition of C_2 and (241) uses the bound in (235). Also note that we used the fact that $\|f\| = M_n = \mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}$ in (242).

Finally in Step 4, by plugging (243) back into (231), we obtain that for all n large enough (depending on δ),

$$\begin{aligned}
&C_1(f_{X_n}(A^n)|E^nX_n|P_{AE}^n \times P_{X_n})\\
&\quad\le \sum_{Q_{AE}\in\mathcal{P}_n(\mathcal{A}\times\mathcal{E}):b_n(Q_{AE})\le L+\delta} P_{AE}^n(\mathcal{T}_{Q_{AE}})\big(L - (1-\delta)b_n(Q_{AE})\big) + O\left(\frac{\log n}{\sqrt{n}}\right). && (244)
\end{aligned}$$

Let $V := V(A|E|P_{AE})$. By the central limit-type convergence in (239), we obtain

$$\limsup_{n\to\infty} \frac{1}{\sqrt{n}} C_1(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n}) \leq \int_{-\infty}^{L+\delta} \frac{L-(1-\delta)b}{\sqrt{2\pi V}} \,\mathrm{e}^{-b^2/(2V)} \,\mathrm{d}b. \tag{245}$$

By a change of variables to $x := b/\sqrt{V}$ and taking $\delta \downarrow 0$, we immediately obtain the direct part (upper bound) of Case (B) in (72).

c) Proof of upper bounds for Case (C): For Case (C), the upper bound to (73) can be obtained by specializing the one-shot bound in (107) with $\epsilon = 1$, $M_n = \mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}$ and $c = \mathrm{e}^{-n^{1/4}}$. With these choices, we have

$$\begin{aligned}
&\mathrm{e}^{-sC_{1-s}(f_{X_n}(A^n)|E^n X_n|P_{AE}^n \times P_{X_n})} \\
&\geq P_{AE}^n\left\{(\mathbf{a},\mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \leq \frac{\mathrm{e}^{-n^{1/4}}}{\mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}}\right\}\left(\frac{1}{1+\mathrm{e}^{-n^{1/4}}}\right)^s && (246)\\
&= P_{AE}^n\left\{(\mathbf{a},\mathbf{e}) : \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[-\log P_{A|E}(a_i|e_i) - H(A|E|P_{AE})\big] \geq L + \frac{1}{\sqrt{n}}\right\}\left(\frac{1}{1+\mathrm{e}^{-n^{1/4}}}\right)^s. && (247)
\end{aligned}$$

The probability is an information spectrum [47] term with n independent and identically distributed random variables and since $P_{A_iE_i} = P_{AE}$ for each $1 \leq i \leq n$,

$$\mathbb{E}_{P_{A_iE_i}}\big[-\log P_{A|E}(A_i|E_i)\big] = H(A|E|P_{AE}), \tag{248}$$

$$\mathsf{Var}_{P_{A_iE_i}}\big[-\log P_{A|E}(A_i|E_i)\big] = V(A|E|P_{AE}). \tag{249}$$

So by the central limit theorem, the right-hand-side of (247) converges uniformly as follows:

$$\lim_{n\to\infty} P_{AE}^n\left\{(\mathbf{a},\mathbf{e}) : \frac{1}{\sqrt{n}}\sum_{i=1}^n \big[-\log P_{A|E}(a_i|e_i) - H(A|E|P_{AE})\big] \geq L + \frac{1}{\sqrt{n}}\right\} = \Phi\Big(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\Big). \tag{250}$$

Plugging (250) into (247), taking the logarithm, and normalizing by $-s$ yields the upper bound to (73).

In a similar way, the upper bound to (74) can be obtained by specializing the one-shot bound in (108) with $\epsilon = 1$, $M_n = \mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}$ and $c = \mathrm{e}^{-n^{1/4}}$. The calculation for the specialization is similar to the converse part which is detailed in full in (280)–(283) in the next section. This completes the proof for the direct part of Case (C) of Theorem 3.

2) Converse Parts:

a) Proof of lower bounds for Case (A): We now prove the lower bounds for Case (A). The first two bounds can be shown using the data processing inequalities in (26)–(28). In particular, the lower bound to (68) can be evaluated as follows:

$$\begin{aligned}
C_{1+s}(f(A^n)|E^n|P_{AE}^n) &= nH_{1+s}(A|E|P_{AE}) + \sqrt{n}L - H_{1+s}(f(A^n)|E^n|P_{AE}^n) && (251)\\
&\geq nH_{1+s}(A|E|P_{AE}) + \sqrt{n}L - H_{1+s}(A^n|E^n|P_{AE}^n) && (252)\\
&= nH_{1+s}(A|E|P_{AE}) + \sqrt{n}L - nH_{1+s}(A|E|P_{AE}) && (253)\\
&= \sqrt{n}L, && (254)
\end{aligned}$$

where (252) follows from (27). The lower bound to (69) follows completely analogously using (28).

The lower bound to (70) can be shown by first relaxing (114) as follows:

$$\begin{aligned}
&\mathrm{e}^{sC_{1+s}(f(A)|E|P_{AE})} \\
&\geq \sum_{(a,e):P_{A|E}(a|e)\geq \frac{c}{M}} P_E(e)P_{A|E}(a|e)^{1+s}M^s + \sum_e P_E(e)P_{A|E=e}\left\{(a,e) : P_{A|E}(a|e) < \frac{c}{M}\right\}^{1+s} && (255)\\
&\geq \sum_e P_E(e)P_{A|E=e}\left\{(a,e) : P_{A|E}(a|e) < \frac{c}{M}\right\}^{1+s} && (256)\\
&\geq P_{AE}\left\{(a,e) : P_{A|E}(a|e) < \frac{c}{M}\right\}^{1+s} && (257)\\
&= \left[1 - P_{AE}\left\{(a,e) : P_{A|E}(a|e) \geq \frac{c}{M}\right\}\right]^{1+s} && (258)\\
&\geq 1 - (1+s)P_{AE}\left\{(a,e) : P_{A|E}(a|e) \geq \frac{c}{M}\right\} && (259)
\end{aligned}$$

where (257) uses Jensen's inequality (for the convex function $t \mapsto t^{1+s}$) and (259) uses the inequality $(1-x)^{1+s} \geq 1-(1+s)x$ (also due to the convexity of $t \mapsto t^{1+s}$). Hence we have for the n-shot setting

$$sC_{1+s}(f(A^n)|E^n|P_{AE}^n) \geq \log\left(1-(1+s)P_{AE}^n\Big\{(\mathbf{a},\mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M_n}\Big\}\right). \tag{260}$$

Applying the modified Gärtner-Ellis theorem derived in Hayashi-Tan [62, Appendix A] to the sequence of random variables $-\log P_{A|E}^n(A^n|E^n)$ with $M_n = \mathrm{e}^{nH_{1+s}(A|E|P_{AE})+\sqrt{n}L}$ and $c=1$, we have

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\log P_{AE}^n\Big\{(\mathbf{a},\mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \geq \frac{c}{M_n}\Big\} = -sL. \tag{261}$$

The modification here is due to the different normalization of $\sqrt{n}$ as opposed to the normalization by n in the usual Gärtner-Ellis theorem in [48]. Also see Remark (a) to Theorem 2.3.6 in [48]. Combining (261) with (260) yields the lower bound to (70). The lower bound to (71) can be proved in a completely analogous way by relaxing the one-shot bound in (116).

b) Proof of lower bound for Case (B): For the converse part of Case (B), we use Theorem 8 of [16], which analyzes the second-order asymptotics of intrinsic randomness [47, Ch. 2] [63]. Define the second-order coding rate at length n as

$$L_n := \frac{1}{\sqrt{n}}\big(\log M_n - nH(A|E|P_{AE})\big) \tag{262}$$

and the distribution function $F_n^{(\mathbf{e})}$ which is dependent on $\mathbf{e}$ as

$$F_n^{(\mathbf{e})}(x) := P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a} : -\frac{1}{n}\log P_{A^n|E^n=\mathbf{e}}(\mathbf{a}) \leq H(A|E|P_{AE}) + \frac{x}{\sqrt{n}}\right\}. \tag{263}$$

Now, from the proof of Theorem 8 of [16] (second column page 4634), we deduce that for each $\mathbf{e} \in \mathcal{E}^n$,

$$\begin{aligned}
&H(f(A^n)|P_{A^n|E^n=\mathbf{e}})\\
&\leq \sqrt{n}\int_{-\infty}^{L_n} a\,\mathrm{d}F_n^{(\mathbf{e})}(a) + nH(A|E|P_{AE})\\
&\quad + P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a} : P_{A^n|E^n=\mathbf{e}}(\mathbf{a}) \leq \frac{1}{M_n}\right\}\left(\sqrt{n}L_n - \log P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a} : P_{A^n|E^n=\mathbf{e}}(\mathbf{a}) \leq \frac{1}{M_n}\right\}\right).
\end{aligned} \tag{264}$$

Now note that $F_n^{(\mathbf{e})}(x)$ depends only on $\mathbf{e}$ through its type. Our next step is to take the expectation of (264) over $\mathbf{e}$ with distribution P_E^n. Let

$$g(\mathbf{e}) := P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a} : P_{A^n|E^n=\mathbf{e}}(\mathbf{a}) \leq \frac{1}{M_n}\right\}. \tag{265}$$

Since $t \mapsto -t\log t$ is concave, by Jensen's inequality, we have

$$\mathbb{E}_{P_{E^n}}[g(E^n)(\gamma - \log g(E^n))] \leq \mathbb{E}_{P_{E^n}}[\gamma g(E^n)] - \mathbb{E}_{P_{E^n}}[g(E^n)]\log\mathbb{E}_{P_{E^n}}[g(E^n)]. \tag{266}$$

Now define the averaged distribution function as

$$F_n(x) := \sum_{\mathbf{e}} P_E^n(\mathbf{e})F_n^{(\mathbf{e})}(x) = P_{AE}^n\left\{(\mathbf{a},\mathbf{e}) : -\frac{1}{n}\log P_{A|E}^n(\mathbf{a}|\mathbf{e}) \leq H(A|E|P_{AE}) + \frac{x}{\sqrt{n}}\right\}. \tag{267}$$

Let $\gamma := \sqrt{n}L_n$. From (266) and the definition of $F_n(x)$,

$$\begin{aligned}
&H(f(A^n)|E^n|P_{AE}^n)\\
&\leq \sqrt{n}\int_{-\infty}^{L_n} a\,\mathrm{d}F_n(a) + nH(A|E|P_{AE})\\
&\quad + P_{AE}^n\left\{(\mathbf{a},\mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \leq \frac{1}{M_n}\right\}\left(\sqrt{n}L_n - \log P_{AE}^n\left\{(\mathbf{a},\mathbf{e}) : P_{A|E}^n(\mathbf{a}|\mathbf{e}) \leq \frac{1}{M_n}\right\}\right).
\end{aligned} \tag{268}$$

Thus, by invoking the definition of L_n in (262) and F_n in (267), we obtain the inequality

$$\frac{1}{\sqrt{n}}\left(H(f(A^n)|E^n|P_{AE}^n)-nH(A|E|P_{AE})\right)$$
$$\leq \int_{-\infty}^{L_n} a\,\mathrm{d}F_n(a)+P_{AE}^n\left\{(\mathbf{a},\mathbf{e}):P_{A|E}^n(\mathbf{a}|\mathbf{e})\leq\frac{1}{M_n}\right\}\left(L_n-\frac{\log P_{AE}^n\left\{(\mathbf{a},\mathbf{e}):P_{A|E}^n(\mathbf{a}|\mathbf{e})\leq\frac{1}{M_n}\right\}}{\sqrt{n}}\right) \tag{269}$$
$$=\int_{-\infty}^{L_n} a\,\mathrm{d}F_n(a)+(1-F_n(L_n))\left(L_n-\frac{\log(1-F_n(L_n))}{\sqrt{n}}\right). \tag{270}$$

By the central limit theorem

$$F_n(x)\to F(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi V}}\mathrm{e}^{-y^2/(2V)}\,\mathrm{d}y,\quad \forall\, x\in\mathbb{R}. \tag{271}$$

Taking the $\limsup$ of (270), and using the central limit result in (271), we obtain

$$\limsup_{n\to\infty}\frac{1}{\sqrt{n}}\left(H(f(A^n)|E^n|P_{AE}^n)-nH(A|E|P_{AE})\right)\leq\int_{-\infty}^{L} a\,\mathrm{d}F(a)+L(1-F(L)). \tag{272}$$

Since, we have the simple relation

$$C_1(f(A^n)|E^n|P_{AE}^n)=D(P_{f(A^n),E^n}\|P_{\mathrm{mix},f(\mathcal{A}^n)}\times P_{E^n}) \tag{273}$$
$$=-H(f(A^n)|E^n|P_{AE}^n)+\log M_n \tag{274}$$
$$=-H(f(A^n)|E^n|P_{AE}^n)+nH(A|E|P_{AE})+\sqrt{n}L, \tag{275}$$

we immediately obtain the desired lower bound for the second-order asymptotics of C_1:

$$\liminf_{n\to\infty}\frac{1}{\sqrt{n}}C_1(f(A^n)|E^n|P_{AE}^n)\geq\int_{-\infty}^{L}(L-a)\,\mathrm{d}F(a)=\int_{-\infty}^{L/\sqrt{V}}\frac{L-\sqrt{V}x}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\,\mathrm{d}x. \tag{276}$$

c) Proof of lower bounds for Case (C): For Case (C), the first part of the maximum in the lower bound in (73), namely $\Gamma_1(s,L)$ in (64), follows from (117) and the second part of the maximum, namely $\Gamma_2(s,L)$ in (65), follows from (112) with the common choice of $c=\mathrm{e}^{n^{1/4}}$. In particular, specializing the bound in one-shot bound in (112) with this choice of c, we obtain

$$\mathrm{e}^{-sC_{1-s}(f(A^n)|E^n|P_{AE}^n)}\leq(\mathrm{e}^{n^{1/4}})^{-s}+P_{AE}^n\left\{(\mathbf{a},\mathbf{e}):P_{A|E}^n(\mathbf{a}|\mathbf{e})\leq\frac{\mathrm{e}^{n^{1/4}}}{\mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}}\right\}^{1-s} \tag{277}$$

where we trivially upper bounded the first probability in the one-shot bound by 1. The first term in (277) goes to zero (since $s>0$) while the second term is an information spectrum term that asymptotically behaves as

$$\lim_{n\to\infty}P_{AE}^n\left\{(\mathbf{a},\mathbf{e}):P_{A|E}^n(\mathbf{a}|\mathbf{e})\leq\frac{\mathrm{e}^{n^{1/4}}}{\mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}}\right\}=\Phi\left(-\frac{L}{\sqrt{V(A|E|P_{AE})}}\right) \tag{278}$$

by the central limit theorem and the statistics computed in (248)–(249). Hence, taking the logarithm in (277), and normalizing by $-s$, we obtain the second term in the maximum in the lower bound in (73), namely $\Gamma_2(s,L)$. In exactly the same way, specializing the bound in (117), we obtain

$$\mathrm{e}^{-sC_{1-s}(f(A^n)|E^n|P_{AE}^n)}\leq(\mathrm{e}^{n^{1/4}})^{-s}+2^{\frac{s}{1-s}}s^{\frac{s}{1-s}}P_{AE}^n\left\{(\mathbf{a},\mathbf{e}):P_{A|E}^n(\mathbf{a}|\mathbf{e})\leq\frac{\mathrm{e}^{n^{1/4}}}{\mathrm{e}^{nH(A|E|P_{AE})+\sqrt{n}L}}\right\}. \tag{279}$$

Applying the central limit theorem to the probability in the second term recovers $\Gamma_1(s,L)$ in the lower bound in (73).

The method to obtain the two terms in the maximum in the lower bound in (74) is more complicated than that for (73) because we need to condition on various sequences $\mathbf{e}\in\mathcal{E}^n$. In particular, to obtain the lower bound $\Psi_1(s,L)$ in (66), we evaluate (110) with $c=\mathrm{e}^{n^{1/4}}$. We obtain

$$\mathrm{e}^{-\frac{s}{1-s}C_{1-s}^{\uparrow}(f(A^n)|E^n|P_{AE}^n)}$$
$$\leq 2^{\frac{s}{1-s}}\left[(\mathrm{e}^{n^{1/4}})^{-\frac{s}{1-s}}+\sum_{\mathbf{e}}P_E^n(\mathbf{e})\left(2^{\frac{s}{1-s}}s^{\frac{s}{1-s}}(1-s)P_{A^n|E^n=\mathbf{e}}\left\{\mathbf{a}:P_{A|E}^n(\mathbf{a}|\mathbf{e})<\frac{c}{M_n}\right\}\right)^{\frac{1}{1-s}}\right]. \tag{280}$$

As usual, the first term goes to zero. To compute the probability in the second term, let us denote the type (empirical distribution) [14] of $\mathbf{e}$ by $Q_{\mathbf{e}} \in \mathcal{P}_n(\mathcal{E})$ for the moment. Then we have

$$\left| P_{A^n|E^n=\mathbf{e}}\left\{ \mathbf{a} : P^n_{A|E}(\mathbf{a}|\mathbf{e}) < \frac{c}{M_n} \right\} - \Phi\left(-\frac{L + H(A|E|P_{AE}) - H(A|E|P_{AE}\|Q_{\mathbf{e}})}{\sqrt{V_2(A|E|P_{AE}\|Q_{\mathbf{e}})}} \right) \right| \le O\left(\frac{1}{\sqrt{n}}\right) \tag{281}$$

by the Berry-Esseen theorem [64, Sec. XVI.7], where the conditional entropy given another distribution $Q_{\mathbf{e}}$, denoted as $H(A|E|P_{AE}\|Q_{\mathbf{e}})$, was defined in (12), and conditional varentropy given another distribution $Q_{\mathbf{e}}$ is defined as

$$V_2(A|E|P_{AE}\|Q_{\mathbf{e}}) := \sum_e Q_{\mathbf{e}}(e) \sum_a P_{A|E}(a|e) \Big[\log \frac{1}{P_{A|E}(a|e)} - H(A|P_{A|E=e}) \Big]^2. \tag{282}$$

Note that $V_2(A|E|P_{AE}\|P_E) = V_2(A|E|P_{AE})$ defined in (62). In (281), the remainder term $O(\frac{1}{\sqrt{n}})$ is uniform in L and $Q_{\mathbf{e}}$. We now plug this into (280) and notice that we are then averaging over all types $Q_{\mathbf{e}}$ (where E^n has distribution P^n_E). Now, employing a weak (expectation) form of the Berry-Esseen theorem [65, Thm. 2.2.14] with $x = H(A|E|P_{AE}) - H(A|E|P_{AE}\|Q_{\mathbf{e}})$ yields

$$\begin{aligned} &\mathrm{e}^{-\frac{s}{1-s} C^{\uparrow}_{1-s}(f(A^n)|E^n|P^n_{AE})} \\ &\quad \le 2^{\frac{s}{1-s}} \Bigg[O\Big(n^{-\frac{s}{4(1-s)}}\Big) + \Big(2^{\frac{s}{1-s}} s^{\frac{s}{1-s}} (1-s)\Big)^{\frac{1}{1-s}} \\ &\qquad \times \int_{-\infty}^{\infty} \Big[\Phi\Big(-\frac{L+x}{\sqrt{V_2(A|E|P_{AE})}} \Big) + O\Big(\frac{1}{\sqrt{n}}\Big) \Big]^{\frac{1}{1-s}} \frac{\mathrm{e}^{-x^2/(2V_1(A|E|P_{AE}))}}{\sqrt{2\pi V_1(A|E|P_{AE})}} \,\mathrm{d}x \Bigg] + O\Big(\frac{1}{\sqrt{n}}\Big). \end{aligned} \tag{283}$$

Now we take the logarithm, divide both sides by $-\frac{s}{1-s}$, and take the limit as $n \to \infty$. This yields the lower bound $\Psi_1(s, L)$ in (66). Note that here unlike in the steps leading to (100), we cannot add V_1 and V_2 due the exponentiation of the first term by $\frac{1}{1-s}$ in the integral.

Using similar techniques, we can obtain the lower bound $\Psi_2(s, L)$ defined in (67) from (113). In particular, evaluate (113) with the same choice of c. Here, in fact, no averaging over E^n is needed because the first term in (113) vanishes by our choice of $c = \mathrm{e}^{n^{1/4}}$. Thus, we obtain the lower bound in (74).

This completes the proof of the converse parts of Theorem 3.

VIII. Conclusion

A. Summary

We have derived the fundamental limits of the asymptotic behavior of the equivocation when a hash function f is applied to the source (Theorem 1). We have also showed that optimal key generation rates change when we use alternative Rényi information measures (Corollary 1). Under these Rényi quantities, we have evaluated the corresponding exponential rates of decay of the security measures (Theorem 2) as well as their second-order coding rates (Theorems 3 and 4). The Rényi information measures generalize the ubiquitous Shannon information measures and may be useful in many settings as described in the Introduction. To establish our asymptotic theorems, we have introduced new families of non-asymptotic achievability and converse bounds on the Rényi information measures and their Gallager counterparts and used various probabilistic limit theorems (such as large deviation theorems and the central limit theorem) to evaluate these bounds when the number of realizations of the joint source tends to infinity.

B. Future Research Directions

In the future, we plan to explore various extensions to the results contained herein.

1) We would like to study security problems such as the *remaining or residual uncertainty* of a source A^n when another party observes a compressed version $f(A^n) \in \mathcal{M} := \{1, \dots, M_n\}$ and another correlated source E^n. Namely, we aim to study the asymptotic behavior of the conditional Rényi entropy $H_{1+s}(A^n|f(A^n), E^n|P^n_{AE})$ and its Gallager counterpart $H^{\uparrow}_{1+s}(A^n|f(A^n), E^n|P^n_{AE})$.
2) Another set of related problems involve the analyses of the asymptotic behavior of $H_{1+s}(f(A^n)|E^n|P^n_{AE})$ and $H^{\uparrow}_{1+s}(f(A^n)|E^n|P^n_{AE})$. These represent the uncertainties of an eavesdropper with regard to the message

index $f(A^n) \in \mathcal{M}$. The eavesdropper, however, is equipped with correlated observations E^n. We anticipate that some of the techniques developed in the current paper may be useful to perform various calculations.

3) We focused primarily on analyzing C_{1+s} and $C_{1+s}^{\uparrow}$ for $s \in [-1, 1]$. It may be of interest to study the various asymptotic behaviors of C_{1+s} and $C_{1+s}^{\uparrow}$ for general $s \in \mathbb{R}$ since for example, $H_{\min} = \lim_{s\to\infty} H_{1+s}$ and $H_{\min}$ [6]–[9] is a fundamental quantity in cryptography and information-theoretic security as mentioned in Section I-A. Indeed, $\mathrm{e}^{-H_{\min}(A|E|P_{AE})}$ is the best (highest) probability of successfully guessing A given E. As remarked after Theorems 1 and 3, we already have the converse parts for all $s \geq 0$ for the results in (47), (48), (68) and (69). They follow immediately from various information processing inequalities. It would be ideal, though challenging, to complete the story.
4) Lastly, we aim to apply the results and techniques derived herein to information-theoretic security problems such as the wiretap channel [12] and secret key agreement [19] as was done by various researchers in [17], [18], [20], [21].

Appendix A
Proof of Lemma 1

A. *Proof of* (102)

Proof: The derivation here is similar to that in [17], [22] for universal$_2$ hash functions. Throughout, for any function $f : \mathcal{A} \to \mathcal{M}$, we let

$$f^{-1}(i) := \{a \in \mathcal{A} : f(a) = i\}, \qquad \forall\, i \in \mathcal{M}. \tag{284}$$

Now, for any a, due to the ϵ-almost universal$_2$ property of f_X, we have

$$\begin{aligned} \mathbb{E}_X \sum_{a' \in f_X^{-1}(f_X(a))} P_{A|E}(a'|e) &\leq P_{A|E}(a|e) + \frac{\epsilon}{M} \sum_{a' \neq a} P_{A|E}(a'|e) && (285)\\ &\leq P_{A|E}(a|e) + \frac{\epsilon}{M}. && (286) \end{aligned}$$

Starting from the definition of the conditional Rényi divergence, we have

$$\begin{aligned} &\mathrm{e}^{-sH_{1+s}(f_X(A)|EX|P_{AE}\times P_X)} \\ &= \mathbb{E}_X \sum_e P_E(e) \sum_{i=1}^{M} \bigg(\sum_{a \in f_X^{-1}(i)} P_{A|E}(a|e) \bigg)^{1+s} && (287)\\ &= \mathbb{E}_X \sum_e P_E(e) \sum_a P_{A|E}(a|e) \bigg(\sum_{a' \in f_X^{-1}(f_X(a))} P_{A|E}(a'|e) \bigg)^{s} && (288)\\ &\leq \sum_e P_E(e) \sum_a P_{A|E}(a|e) \bigg(\mathbb{E}_X \sum_{a' \in f_X^{-1}(f_X(a))} P_{A|E}(a'|e) \bigg)^{s} && (289)\\ &\leq \sum_e P_E(e) \sum_a P_{A|E}(a|e) \bigg(P_{A|E}(a|e) + \frac{\epsilon}{M} \bigg)^{s} && (290)\\ &\leq \sum_e P_E(e) \sum_a P_{A|E}(a|e) \bigg(P_{A|E}(a|e)^s + \Big(\frac{\epsilon}{M}\Big)^{s} \bigg) && (291)\\ &= \Big(\frac{\epsilon}{M}\Big)^{s} + \sum_e P_E(e) \sum_a P_{A|E}(a|e)^{1+s} && (292)\\ &= \frac{\epsilon^s}{M^s} + \mathrm{e}^{-sH_{1+s}(A|E|P_{AE})}, && (293) \end{aligned}$$

where (289), we used the concavity of $t \mapsto t^s$ for $s \in [0, 1]$, in (290) we used the fact that f_X is a ϵ-almost universal$_2$ hash function, and in (291) we used the inequality $(\sum_i a_i)^s \leq \sum_i a_i^s$ for $s \in [0, 1]$ [13, Problem 4.15(f)].

We remark that the sequence of steps in (287) to (293) is inspired by the work of Hayashi [17] who derived a similar result but for Shannon-type quantities instead of Rényi-type quantities as we do here.

By (34), we have

$$C_{1+s}(f_X(A)|EX|P_{AE}\times P_X)=\log M-H_{1+s}(f_X(A)|EX|P_{AE}\times P_X). \tag{294}$$

Uniting (293) and (294) proves (102) as desired. ■

B. Proof of (103)

Proof: Along exactly the same lines, we also have

$$\begin{aligned}
&\mathrm{e}^{-\frac{s}{1+s}H^{\uparrow}_{1+s}(f_X(A)|EX|P_{AE}\times P_X)}\\
&=\mathbb{E}_X\sum_e P_E(e)\Big(\sum_{i=1}^{M}\Big(\sum_{a\in f_X^{-1}(i)}P_{A|E}(a|e)\Big)^{1+s}\Big)^{\frac{1}{1+s}} && (295)\\
&=\mathbb{E}_X\sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\Big(\sum_{a'\in f_X^{-1}(f_X(a))}P_{A|E}(a'|e)\Big)^{s}\Big)^{\frac{1}{1+s}} && (296)\\
&\le\sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\Big(\mathbb{E}_X\sum_{a'\in f_X^{-1}(f_X(a))}P_{A|E}(a'|e)\Big)^{s}\Big)^{\frac{1}{1+s}} && (297)\\
&\le\sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\Big(P_{A|E}(a|e)+\frac{\epsilon}{M}\Big)^{s}\Big)^{\frac{1}{1+s}} && (298)\\
&\le\sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\Big(P_{A|E}(a|e)^s+\Big(\frac{\epsilon}{M}\Big)^{s}\Big)\Big)^{\frac{1}{1+s}} && (299)\\
&=\sum_e P_E(e)\Big(\Big(\frac{\epsilon}{M}\Big)^{s}+\sum_a P_{A|E}(a|e)^{1+s}\Big)^{\frac{1}{1+s}} && (300)\\
&\le\sum_e P_E(e)\Big(\Big(\frac{\epsilon}{M}\Big)^{\frac{s}{1+s}}+\Big(\sum_a P_{A|E}(a|e)^{1+s}\Big)^{\frac{1}{1+s}}\Big) && (301)\\
&=\Big(\frac{\epsilon}{M}\Big)^{\frac{s}{1+s}}+\sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)^{1+s}\Big)^{\frac{1}{1+s}} && (302)\\
&=\frac{\epsilon^{\frac{s}{1+s}}}{M^{\frac{s}{1+s}}}+\mathrm{e}^{-\frac{s}{1+s}H^{\uparrow}_{1+s}(A|E|P_{AE})}. && (303)
\end{aligned}$$

Combining this with the relation between $H^{\uparrow}_{1+s}$ and $C^{\uparrow}_{1+s}$ in (36), we obtain (103). ■

C. Proof of (104)

Proof: For any a, we have

$$\begin{aligned}
&\mathbb{E}_X\sum_{a'\in f_X^{-1}(f_X(a))}P_{A|E}(a'|e)\\
&\le P_{A|E}(a|e)+\frac{\epsilon}{M}\sum_{a'\neq a}P_{A|E}(a'|e) && (304)\\
&\le P_{A|E}(a|e)+\frac{\epsilon}{M} && (305)\\
&\le 2\max\Big\{P_{A|E}(a|e),\frac{\epsilon}{M}\Big\}. && (306)
\end{aligned}$$

First we observe that when $P_{A|E}(a|e)\le\frac{c}{M}$, we have

$$\mathbb{E}_X\sum_{a'\in f_X^{-1}(f_X(a))}P_{A|E}(a'|e)\le P_{A|E}(a|e)+\frac{\epsilon}{M}\le\frac{c+\epsilon}{M}. \tag{307}$$

Now we have

$$
\begin{aligned}
& e^{-sC_{1-s}(f_X(A)|EX|P_{AE}\times P_X)} \\
&= \frac{1}{M^s} e^{sH_{1-s}(f_X(A)|EX|P_{AE}\times P_X)} && (308)\\
&\ge \frac{1}{M^s}\sum_e P_E(e)\sum_a P_{A|E}(a|e)\Big(\mathbb{E}_X \sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s} && (309)\\
&\ge \sum_{a,e} P_{AE}(a,e)\big(2\max\{MP_{A|E}(a|e),\epsilon\}\big)^{-s} && (310)\\
&\ge 2^{-s}\sum_{a,e} P_{AE}(a,e)\min\{P_{A|E}(a|e)^{-s}M^{-s},\epsilon^{-s}\} && (311)\\
&= 2^{-s}\sum_{a,e:P_{A|E}(a|e)\ge \epsilon M^{-1}} P_{AE}(a,e)P_{A|E}(a|e)^{-s}M^{-s} \\
&\quad + 2^{-s}\sum_{a,e:P_{A|E}(a|e)<\epsilon M^{-1}} P_{AE}(a,e)\epsilon^{-s}, && (312)
\end{aligned}
$$

where in (309) we used the convexity of $x\mapsto x^{-s}$ where $s\in[0,1]$ and $x\ge 0$ and in (310), we used (306), Thus, we obtain (104). ■

D. *Proof of* (105)

Proof: Using (306) and the convexity of $a\mapsto a^{\frac{1}{1-s}}$ we have

$$
\begin{aligned}
& e^{\frac{s}{1-s}H^{\uparrow}_{1-s}(f_X(A)|EX|P_{AE}\times P_X)} \\
&= \mathbb{E}_X\sum_e P_E(e)\Big(\sum_{i=1}^{M}\Big(\sum_{a\in f_X^{-1}(i)} P_{A|E}(a|e)\Big)^{1-s}\Big)^{\frac{1}{1-s}} && (313)\\
&= \mathbb{E}_X\sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\Big(\sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s}\Big)^{\frac{1}{1-s}} && (314)\\
&\ge \sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\Big(\mathbb{E}_X\sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s}\Big)^{\frac{1}{1-s}} && (315)\\
&\ge \sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\Big(P_{A|E}(a|e)+\frac{\epsilon}{M}\Big)^{-s}\Big)^{\frac{1}{1-s}} && (316)\\
&\ge \sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\Big(2\max\Big\{P_{A|E}(a|e),\frac{\epsilon}{M}\Big\}\Big)^{-s}\Big)^{\frac{1}{1-s}} && (317)\\
&= 2^{-\frac{s}{1-s}}\sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)\min\Big\{P_{A|E}(a|e)^{-s},\frac{\epsilon^{-s}}{M^{-s}}\Big\}\Big)^{\frac{1}{1-s}} && (318)\\
&= 2^{-\frac{s}{1-s}}\sum_e P_E(e)\Big(\sum_{a:P_{A|E}(a|e)\ge\frac{\epsilon}{M}} P_{A|E}(a|e)^{1-s}+\frac{\epsilon^{-s}}{M^{-s}}\sum_{a:P_{A|E}(a|e)<\frac{\epsilon}{M}} P_{A|E}(a|e)\Big)^{\frac{1}{1-s}} && (319)\\
&\ge 2^{-\frac{s}{1-s}}\sum_e P_E(e)\Big(\Big(\sum_{a:P_{A|E}(a|e)\ge\frac{\epsilon}{M}} P_{A|E}(a|e)\Big)^{\frac{1}{1-s}}+\Big(\frac{\epsilon^{-s}}{M^{-s}}\sum_{a:P_{A|E}(a|e)<\frac{\epsilon}{M}} P_{A|E}(a|e)^{1-s}\Big)^{\frac{1}{1-s}}\Big) && (320)\\
&\ge 2^{-\frac{s}{1-s}}\sum_e P_E(e)\Big(\Big(\sum_{a:P_{A|E}(a|e)\ge\frac{\epsilon}{M}} P_{A|E}(a|e)^{1-s}\Big)^{\frac{1}{1-s}}+\Big(\frac{\epsilon^{-s}}{M^{-s}}\sum_{a:P_{A|E}(a|e)<\frac{\epsilon}{M}} P_{A|E}(a|e)\Big)^{\frac{1}{1-s}}\Big) && (321)
\end{aligned}
$$

$$= 2^{-\frac{s}{1-s}} \sum_e P_E(e) \Bigg(\sum_{a: P_{A|E}(a|e) \ge \frac{\epsilon}{M}} P_{A|E}(a|e)^{1-s} \Bigg)^{\frac{1}{1-s}}$$

$$+ 2^{-\frac{s}{1-s}} \frac{M^{\frac{s}{1-s}}}{\epsilon^{\frac{s}{1-s}}} \sum_e P_E(e) \Bigg(\sum_{a: P_{A|E}(a|e) < \frac{\epsilon}{M}} P_{A|E}(a|e) \Bigg)^{\frac{1}{1-s}}. \tag{322}$$

Thus we obtain (105). ■

APPENDIX B
PROOF OF LEMMA 2

A. *Proof of* (106)

Proof: Since $(1+x)^{\frac{1}{1+s}} \le 1 + \frac{1}{1+s}x$, and $x \mapsto x^{\frac{1}{1+s}}$ is concave for $s \in [0,1]$, we have

$$M^{\frac{s}{1+s}} \mathrm{e}^{-\frac{s}{1+s} H^{\uparrow}_{1+s}(f_X(A)|EX|P_{AE} \times P_X)}$$

$$= M^{\frac{s}{1+s}} \mathbb{E}_X \sum_e P_E(e) \Bigg(\sum_{i=1}^{M} \Big(\sum_{a \in f_X^{-1}(i)} P_{A|E}(a|e) \Big)^{1+s} \Bigg)^{\frac{1}{1+s}} \tag{323}$$

$$= M^{\frac{s}{1+s}} \mathbb{E}_X \sum_e P_E(e) \Bigg(\sum_a P_{A|E}(a|e) \Big(\sum_{a' \in f_X^{-1}(f_X(a))} P_{A|E}(a'|e) \Big)^{s} \Bigg)^{\frac{1}{1+s}} \tag{324}$$

$$\le M^{\frac{s}{1+s}} \sum_e P_E(e) \Bigg(\sum_a P_{A|E}(a|e) \Big(\mathbb{E}_X \sum_{a' \in f_X^{-1}(f_X(a))} P_{A|E}(a'|e) \Big)^{s} \Bigg)^{\frac{1}{1+s}} \tag{325}$$

$$\le M^{\frac{s}{1+s}} \sum_e P_E(e) \Bigg(\sum_a P_{A|E}(a|e) \Big(P_{A|E}(a|e) + \frac{1}{M} \Big)^{s} \Bigg)^{\frac{1}{1+s}} \tag{326}$$

$$\le M^{\frac{s}{1+s}} \sum_e P_E(e) \Bigg(\sum_a P_{A|E}(a|e) \Big(P_{A|E}(a|e)^s + \frac{1}{M^s} \Big) \Bigg)^{\frac{1}{1+s}} \tag{327}$$

$$= M^{\frac{s}{1+s}} \sum_e P_E(e) \Bigg(\frac{1}{M^s} + \sum_a P_{A|E}(a|e)^{1+s} \Bigg)^{\frac{1}{1+s}} \tag{328}$$

$$= \sum_e P_E(e) \Bigg(1 + M^s \sum_a P_{A|E}(a|e)^{1+s} \Bigg)^{\frac{1}{1+s}} \tag{329}$$

$$\le \sum_e P_E(e) \Bigg(1 + \frac{1}{1+s} M^s \sum_a P_{A|E}(a|e)^{1+s} \Bigg) \tag{330}$$

$$= 1 + \frac{1}{1+s} M^s \sum_e P_E(e) \sum_a P_{A|E}(a|e)^{1+s} \tag{331}$$

$$= 1 + \frac{1}{1+s} M^s \mathrm{e}^{-s H_{1+s}(A|E|P_{AE})}. \tag{332}$$

Using (294), we obtain (106). ■

APPENDIX C
PROOF OF LEMMA 3

A. *Proof of* (107)

Proof: Using (307), we have

$$
\begin{aligned}
& e^{-sC_{1-s}(f_X(A)|EX|P_{AE}\times P_X)} \\
&= \frac{1}{M^s} e^{sH_{1-s}(f_X(A)|EX|P_{AE}\times P_X)} && (333)\\
&= \frac{1}{M^s}\mathbb{E}_X \sum_e P_E(e) \sum_{i=1}^{M} \Big(\sum_{a\in f_X^{-1}(i)} P_{A|E}(a|e)\Big)^{1-s} && (334)\\
&= \frac{1}{M^s}\mathbb{E}_X \sum_e P_E(e) \sum_a P_{A|E}(a|e)\Big(\sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s} && (335)\\
&\geq \frac{1}{M^s} \sum_e P_E(e) \sum_a P_{A|E}(a|e)\Big(\mathbb{E}_X \sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s} && (336)\\
&\geq \frac{1}{M^s} \sum_e P_E(e) \sum_{a:P_{A|E}(a|e)\leq \frac{c}{M}} P_{A|E}(a|e)\Big(\mathbb{E}_X \sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s} && (337)\\
&\geq \frac{1}{M^s} \sum_e P_E(e) \sum_{a:P_{A|E}(a|e)\leq \frac{c}{M}} P_{A|E}(a|e)\Big(\frac{c+\epsilon}{M}\Big)^{-s} && (338)\\
&= P_{AE}\Big\{(a,e) : P_{A|E}(a|e) \leq \frac{c}{M}\Big\}\Big(\frac{1}{c+\epsilon}\Big)^{s}. && (339)
\end{aligned}
$$

We obtain (107) as desired. ■

B. *Proof of* (108)

Proof: Using (307), we have

$$
\begin{aligned}
& e^{\frac{s}{1-s}H^{\uparrow}_{1-s}(f_X(A)|EX|P_{AE}\times P_X)} \\
&= \mathbb{E}_X \sum_e P_E(e) \Big(\sum_{i=1}^{M} \Big(\sum_{a\in f_X^{-1}(i)} P_{A|E}(a|e)\Big)^{1-s}\Big)^{\frac{1}{1-s}} && (340)\\
&= \mathbb{E}_X \sum_e P_E(e) \Big(\sum_a P_{A|E}(a|e)\Big(\sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s}\Big)^{\frac{1}{1-s}} && (341)\\
&\geq \sum_e P_E(e) \Big(\sum_a P_{A|E}(a|e)\Big(\mathbb{E}_X \sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s}\Big)^{\frac{1}{1-s}} && (342)\\
&\geq \sum_e P_E(e) \Big(\sum_{a:P_{A|E}(a|e)\leq \frac{c}{M}} P_{A|E}(a|e)\Big(\mathbb{E}_X \sum_{a'\in f_X^{-1}(f_X(a))} P_{A|E}(a'|e)\Big)^{-s}\Big)^{\frac{1}{1-s}} && (343)\\
&\geq \sum_e P_E(e) \Big(\sum_{a:P_{A|E}(a|e)\leq \frac{c}{M}} P_{A|E}(a|e)\Big(\frac{c+\epsilon}{M}\Big)^{-s}\Big)^{\frac{1}{1-s}} && (344)\\
&= \Big(\frac{c+\epsilon}{M}\Big)^{-\frac{s}{1-s}} \sum_e P_E(e) \Big(\sum_{a:P_{A|E}(a|e)\leq \frac{c}{M}} P_{A|E}(a|e)\Big)^{\frac{1}{1-s}}. && (345)
\end{aligned}
$$

By combining with (36), we obtain (108). ■

Appendix D
Proof of Lemma 4

A. *Proof of* (109)

Proof: Define the functions

$$g_1(x,y) := x + y - 2x^{1-s}y^s \tag{346}$$
$$g_2(x,y) := x - 2x^{1-s}y^s. \tag{347}$$

Then, we can show that

$$\min_y g_1(x,y) = x(1 - 2^{\frac{1}{1-s}} s^{\frac{1}{1-s}-1}(1-s)), \tag{348}$$

which is attained when $y = x(2s)^{\frac{1}{1-s}}$. We also define

$$h_{e,1}(m) := \sum_{a \in f^{-1}(m): P_{A|E}(a|e) < \frac{c}{M}} P_{A|E}(a|e) \tag{349}$$
$$h_{e,2}(m) := \sum_{a \in f^{-1}(m): P_{A|E}(a|e) \geq \frac{c}{M}} P_{A|E}(a|e). \tag{350}$$

Hence,

$$P_{f(A)|E}(m|e)^{1-s} = (h_{e,1}(m) + h_{e,2}(m))^{1-s} \leq h_{e,1}(m)^{1-s} + h_{e,2}(m)^{1-s}, \tag{351}$$

which implies from the definitions of g_1 and g_2 that

$$g_1\Big(P_{f(A)|E}(m|e), \frac{1}{M}\Big) \geq g_2\Big(h_{e,2}(m), \frac{1}{M}\Big) + g_1\Big(h_{e,1}(m), \frac{1}{M}\Big). \tag{352}$$

Also, we have

$$h_{e,2}(m)^{1-s} \leq \sum_{a \in f^{-1}(m)} P_{A|E}(a|e)^{1-s}. \tag{353}$$

Thus,

$$
\begin{aligned}
&1 - 2\mathrm{e}^{-sC_{1-s}(f(A)|E|P_{AE})} \\
&= \sum_e P_E(e) \sum_m g_1\Big(P_{f(A)|E}(m|e), \frac{1}{M}\Big) &(354)\\
&\geq \sum_e P_E(e) \bigg(\sum_m g_2\Big(h_{e,2}(m), \frac{1}{M}\Big) + \sum_m g_1\Big(h_{e,1}(m), \frac{1}{M}\Big)\bigg) &(355)\\
&\geq \sum_e P_E(e) \bigg(\sum_m (h_{e,2}(m) - 2h_{e,2}(m)^{1-s}M^s) + \sum_m h_{e,1}(m)(1 - 2^{\frac{1}{1-s}} s^{\frac{s}{1-s}}(1-s))\bigg) &(356)\\
&= 1 - 2c^{-s} \sum_e P_E(e) \sum_m h_{e,2}(m)^{1-s}M^s - 2^{\frac{1}{1-s}} s^{\frac{s}{1-s}}(1-s) \sum_m h_{e,1}(m) &(357)\\
&\geq 1 - 2c^{-s} \sum_e P_E(e) \sum_{a:P_{A|E}(a|e) \geq \frac{c}{M}} P_{A|E}(a|e)^{1-s}M^s - 2^{\frac{1}{1-s}} s^{\frac{s}{1-s}}(1-s) \sum_{a:P_{A|E}(a|e) < \frac{c}{M}} P_{A|E}(a|e) &(358)\\
&= 1 - 2c^{-s} \sum_e P_E(e) \sum_{a:P_{A|E}(a|e) \geq \frac{c}{M}} P_{A|E}(a|e)^{1-s}M^s \\
&\qquad - 2 \cdot 2^{\frac{s}{1-s}} s^{\frac{s}{1-s}}(1-s) P_{AE}\Big\{(a,e) : P_{A|E}(a|e) < \frac{c}{M}\Big\}, &(359)
\end{aligned}
$$

where (355), (356), and (358) follow from (352), (348), and (353) respectively. Hence, we obtain (109). ■

B. *Proof of* (110)

We first state a useful and easy lemma:

Lemma 7. *Let $x, y \geq 0$ and $t \geq 1$. Then we have*

$$(x+y)^t \leq 2^{t-1}(x^t + y^t). \tag{360}$$

Proof: It is clear that $a \mapsto a^t$ is convex for $a \geq 0$. Thus,

$$(x+y)^t = 2^t\Big(\frac{x}{2}+\frac{y}{2}\Big)^t \leq 2^t\Big(\frac{x^t}{2}+\frac{y^t}{2}\Big) = 2^{t-1}(x^t+y^t), \tag{361}$$

which proves the claim. ∎

Proof of (110)*:* Using the previously proved bound in (109) with $|\mathcal{E}| = 1$, we obtain

$$\begin{aligned} \mathrm{e}^{-sC^{\uparrow}_{1-s}(f(A)|P_{A|E=e})} \leq c^{-s} &\sum_{a: P_{A|E}(a|e) \geq \frac{c}{M}} P_{A|E}(a|e)^{1-s} M^{-s} \\ &+ 2^{\frac{s}{1-s}} s^{\frac{s}{1-s}} (1-s) P_{AE}\Big\{(a,e) : P_{A|E}(a|e) \leq \frac{c}{M}\Big\}. \end{aligned} \tag{362}$$

Taking average over P_E and using the bound in (362), we have

$$\begin{aligned} &\mathrm{e}^{-\frac{s}{1-s}C^{\uparrow}_{1-s}(f(A)|E|P_{AE})} \\ &= \sum_e P_E(e)\left(\mathrm{e}^{-sC^{\uparrow}_{1-s}(f(A)|P_{A|E=e})}\right)^{\frac{1}{1-s}} &(363)\\ &\leq \sum_e P_E(e)\Big[c^{-s}\sum_{a: P_{A|E}(a|e) \geq \frac{c}{M}} P_{A|E}(a|e)^{1-s}M^{-s} \\ &\qquad + 2^{\frac{s}{1-s}} s^{\frac{s}{1-s}}(1-s)P_{AE}\Big\{(a,e) : P_{A|E}(a|e) \leq \frac{c}{M}\Big\}\Big]^{\frac{1}{1-s}} &(364)\\ &\leq 2^{\frac{s}{1-s}}\sum_e P_E(e)\Big[\Big(c^{-s}\sum_{a: P_{A|E}(a|e) \geq \frac{c}{M}} P_{A|E}(a|e)^{1-s}M^{-s}\Big)^{\frac{1}{1-s}} \\ &\qquad + \Big(2^{\frac{s}{1-s}} s^{\frac{s}{1-s}}(1-s)P_{AE}\Big\{(a,e) : P_{A|E}(a|e) \leq \frac{c}{M}\Big\}\Big)^{\frac{1}{1-s}}\Big] &(365) \end{aligned}$$

where in the last step, we applied Lemma 7 with $t = \frac{1}{1-s} \geq 1$ to the term in parentheses in (364). Thus we obtain (110). ∎

Appendix E
Proof of Lemma 5

The inequalities in Lemma 5 can be shown by the information processing inequality for Rényi divergence in (11).

A. *Proofs of* (111) *and* (112)

Proof: For every $e \in \mathcal{E}$, define the function $f_e : \mathcal{A} \to \mathcal{M}$ to be

$$f_e := \arg\min_f D_{1-s}(P_{f(A)|E=e} \| P_{\mathrm{mix},\mathcal{M}}). \tag{366}$$

We start with a claim that will be proved at the end of this subsection.

Lemma 8. *For every a such that $P_{A|E=e}(a) \geq \frac{1}{M}$, we have $|f_e^{-1}(f_e(a))| = 1$.*

Now, we partition the set $\mathcal{M}$ into two subsets as follows: $\mathcal{M}_1 := f_e(\{a : P_{A|E=e}(a) \geq \frac{c}{M}\})$ and $\mathcal{M}_2 := \mathcal{M}_1^c$. See Fig. 9. Next, we define the map $g : \mathcal{M} \to \mathcal{M}_1 \cup \{0\}$ as

$$g(i) := \begin{cases} i & \text{if } i \in \mathcal{M}_1 \\ 0 & \text{if } i \in \mathcal{M}_2 \end{cases}. \tag{367}$$

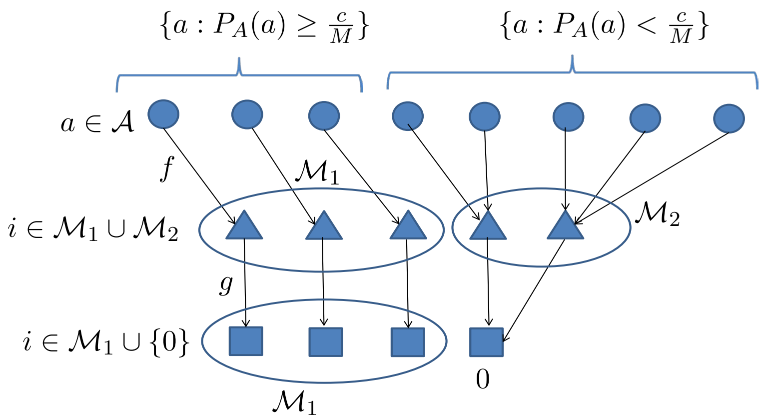

Fig. 9. Illustration of the steps in (368) to (372) where the dependences on $e \in \mathcal{E}$ are suppressed.

Note that the map g depends on $e \in \mathcal{E}$ (through $\mathcal{M}_1$ and $\mathcal{M}_2$) but we suppress this dependence for brevity. Let $P_{f_e(A)|E=e} \circ g^{-1}$ be the "output distribution" taking values on $\mathcal{M}_1 \cup \{0\}$ induced by the "input distribution" $P_{f_e(A)|E=e}$ and the "deterministic channel" g, i.e., for every $i \in \mathcal{M}_1 \cup \{0\}$, $(P_{f_e(A)|E=e} \circ g^{-1})(i) := P_{f_e(A)|E=e}(g^{-1}(i)) = \sum_{j \in \mathcal{M}: g(i)=j} P_{f_e(A)|E=e}(j)$. We also use the notation $P_{\mathrm{mix},\mathcal{M}} \circ g^{-1}$ analogously. Due to the information processing inequality for the Rényi divergence in (11), we obtain

$$
\begin{aligned}
&\mathrm{e}^{-sD_{1-s}(P_{f_e(A)|E=e} \| P_{\mathrm{mix},\mathcal{M}})} \\
&\le \mathrm{e}^{-sD_{1-s}(P_{f_e(A)|E=e} \circ g^{-1} \| P_{\mathrm{mix},\mathcal{M}} \circ g^{-1})} && (368) \\
&= \sum_{i \in \{0\} \cup \mathcal{M}_1} \big((P_{f_e(A)|E=e} \circ g^{-1})(i)\big)^{1-s} \big((P_{\mathrm{mix},\mathcal{M}} \circ g^{-1})(i)\big)^{s} && (369) \\
&= \big((P_{f_e(A)|E=e} \circ g^{-1})(0)\big)^{1-s} \big((P_{\mathrm{mix},\mathcal{M}} \circ g^{-1})(0)\big)^{s} + \sum_{i \in \mathcal{M}_1} \big((P_{f_e(A)|E=e} \circ g^{-1})(i)\big)^{1-s} \Big(\frac{1}{M}\Big)^{s} && (370) \\
&\le P_{A|E=e}\Big\{a : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1-s} + \sum_{i \in \mathcal{M}_1} \big((P_{f_e(A)|E=e} \circ g^{-1})(i)\big)^{1-s} \Big(\frac{1}{M}\Big)^{s} && (371) \\
&= P_{A|E=e}\Big\{a : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1-s} + \sum_{a : P_{A|E=e}(a) \ge \frac{c}{M}} P_{A|E}(a|e)^{1-s} M^{-s}, && (372)
\end{aligned}
$$

where in (369) we used the definition of D_{1-s}, in (370) we split the resulting sum into $\{0\}$ and $\mathcal{M}_1$, in (371) we upper bounded $(P_{\mathrm{mix},\mathcal{M}} \circ g^{-1})(0)$ by 1 and we noted that all the symbols $i \in \mathcal{M}_2 = f_e(\{a : P_{A|E=e}(a) < \frac{c}{M}\})$ are merged into the symbol 0 under g and finally in (372), we used the fact that g and f_e are one-to-one restricted to $\mathcal{M}_1$ and $\{a : P_{A|E=e}(a) \ge \frac{c}{M}\}$ respectively (Lemma 8 and the fact that $c \ge 1$). See Fig. 9 for an illustration of these steps. Taking the average of (372) over $P_E(e)$, we obtain (111).

Furthermore, we have

$$
\begin{aligned}
&\sum_{a : P_{A|E=e}(a) \ge \frac{c}{M}} P_{A|E}(a|e)^{1-s} M^{-s} \\
&\le \sum_{a : P_{A|E=e}(a) \ge \frac{c}{M}} P_{A|E}(a|e) \Big(\frac{c}{M}\Big)^{-s} M^{-s} && (373) \\
&= P_{A|E=e}\Big\{a : P_{A|E=e}(a) \ge \frac{c}{M}\Big\} c^{-s}. && (374)
\end{aligned}
$$

Substituting (374) into the second term in (372) and then taking the average over $P_E(e)$, we obtain (112). ∎

It remains to prove Lemma 8.

Proof of Lemma 8: It suffices to consider the case $|\mathcal{E}| = 1$. Dropping the dependences on e, we denote f_e as f and $P_{A|E=e}$ as P_A in the sequel.

We proceed by contradiction. The essential idea is that an optimal f (given by (366)) must induce a distribution $P_{f(A)}$ on $\mathcal{M}$ that is "as close to uniform as possible" since we are minimizing $D_{1-s}(P_{f(A)}\|P_{\mathrm{mix},\mathcal{M}})$.

Formally, assume, to the contrary, that $|f^{-1}(f(a_1))| \geq 2$ for some $a_1 \in \mathcal{A}$ with $P_A(a_1) \geq \frac{1}{M}$. Because $|f^{-1}(f(a_1))| \geq 2$, there exists $a_2 \neq a_1$ such that $f(a_1) = f(a_2) = j$ for some $j \in \mathcal{M}$. Because $P_A(a_2) > 0$, we have $\sum_{a\in f^{-1}(j)} P_A(a) > \frac{1}{M}$. This in turn implies that there exists $i \in \mathcal{M}$ such that $\sum_{a\in f^{-1}(i)} P_A(a) < \frac{1}{M}$. Recall that f was designed to minimize

$$D_{1-s}(P_{f(A)}\|P_{\mathrm{mix},\mathcal{M}}) = -\frac{1}{s}\log\sum_{i\in\mathcal{M}}\Big(\sum_{a\in f^{-1}(i)} P_A(a)\Big)^{1-s}\Big(\frac{1}{M}\Big)^{s}, \tag{375}$$

or equivalently, to maximize $\sum_i \big(\sum_{a\in f^{-1}(i)} P_A(a)\big)^{1-s}$. Now we create a new hash function

$$\tilde{f}(a) := \begin{cases} f(a) & a \neq a_2 \\ i & a = a_2 \end{cases}. \tag{376}$$

Let u, v, u', v' be any four non-negative numbers such that $u + v = u' + v' = t$ and $p = u/t$, $p' = u'/t$ and $|p - 1/2| < |p' - 1/2|$. This means that $(p, 1-p)$ is closer to the uniform Bernoulli distribution compared to $(p', 1-p')$. Then it is easy to check that

$$(u')^{1-s} + (v')^{1-s} < u^{1-s} + v^{1-s}. \tag{377}$$

Now denoting $\frac{1}{M}+\delta_1 := \sum_{a\in f^{-1}(j)} P_A(a)$, $\frac{1}{M}-\delta_2 := \sum_{a\in f^{-1}(i)} P_A(a)$ for positive numbers δ_1 and δ_2 (positive by the above construction of the sets $f^{-1}(i)$ and $f^{-1}(j)$), and letting $q := P_A(a_2) \leq \delta_1$, $t := (\frac{1}{M}+\delta_1)+(\frac{1}{M}-\delta_2)$, we find that $\frac{1}{t}(\frac{1}{M}+\delta_1-q, \frac{1}{M}-\delta_2+q)$ is closer to the uniform Bernoulli distribution compared to $\frac{1}{t}(\frac{1}{M}+\delta_1, \frac{1}{M}-\delta_2)$. Using inequality (377), we find that

$$\begin{aligned}&\Big(\sum_{a\in f^{-1}(j)} P_A(a)\Big)^{1-s} + \Big(\sum_{a\in f^{-1}(i)} P_A(a)\Big)^{1-s} \\ &\quad< \Big(\sum_{a\in f^{-1}(j)\setminus\{a_2\}} P_A(a)\Big)^{1-s} + \Big(\Big(\sum_{a\in f^{-1}(i)} P_A(a)\Big) + P_A(a_2)\Big)^{1-s}.\end{aligned} \tag{378}$$

Since $f^{-1}(k) = \tilde{f}^{-1}(k)$ for all $k \notin \{i, j\}$, (378) implies that

$$\sum_i\Big(\sum_{a\in f^{-1}(i)} P_A(a)\Big)^{1-s} < \sum_i\Big(\sum_{a\in \tilde{f}^{-1}(i)} P_A(a)\Big)^{1-s}, \tag{379}$$

contradicting the optimality of f. ■

B. *Proof of* (113)

Proof: Here, we employ the following expression for $C^{\uparrow}_{1-s}(A|E|P_{AE})$:

$$\mathrm{e}^{-\frac{s}{1-s}C^{\uparrow}_{1-s}(A|E|P_{AE})} = \frac{1}{|\mathcal{A}|^{\frac{s}{1-s}}}\sum_e P_E(e)\Big(\sum_a P_{A|E}(a|e)^{1-s}\Big)^{\frac{1}{1-s}}. \tag{380}$$

To minimize $\mathrm{e}^{-\frac{s}{1-s}C^{\uparrow}_{1-s}(f(A)|E|P_{AE})}$, it is enough to minimize $\sum_{i\in\mathcal{M}}(\sum_{a\in f^{-1}(i)} P_{A|E=e}(a))^{1-s}$ for each e.

Fortunately, the discussion in the proof in Appendix E-A (and, in particular, the bound (111)) shows that this value is upper bounded by

$$\sum_{a:P_{A|E=e}(a)\geq\frac{c}{M}} P_{A|E}(a|e)^{1-s} + P_{A|E=e}\Big\{a : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1-s} M^s. \tag{381}$$

Thus,

$$
\begin{aligned}
&\mathrm{e}^{-\frac{s}{1-s}C^{\uparrow}_{1-s}(f(A)|E|P_{AE})} && (382)\\
&\le \frac{1}{M^{\frac{s}{1-s}}}\sum_e P_E(e)\bigg(P_{A|E=e}\Big\{a : P_{A|E=e}(a) \ge \frac{c}{M}\Big\}c^{-s}M^s \\
&\qquad + P_{A|E=e}\Big\{a : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1-s}M^s\bigg)^{\frac{1}{1-s}} && (383)\\
&= \sum_e P_E(e)\bigg(P_{A|E=e}\Big\{a : P_{A|E=e}(a) \ge \frac{c}{M}\Big\}c^{-s} \\
&\qquad + P_{A|E=e}\Big\{a : P_{A|E}(a|e) < \frac{c}{M}\Big\}^{1-s}\bigg)^{\frac{1}{1-s}}. && (384)
\end{aligned}
$$

Hence, we obtain (113) as desired. ■

C. *Proofs of* (114) *and* (115)

Proof: The proofs of these bounds are similar to those of (111) and (112) in Appendix E-A and thus are omitted. ■

D. *Proof of* (116)

Proof: The proof of this bound is similar to that of (113) in Appendix E-B and is thus omitted. ■

APPENDIX F
PROOF OF LEMMA 6

A. *Proof of* (117)

Proof: The proof of this bound is similar to the proof of (109) which is presented in Appendix D-A. We provide the details here.

When $P_{A|E}(a|e) < \frac{c}{M}$, we have $P_{A|E}(a|e)^{1-s}M^s \le P_{A|E}(a|e)c^{-s}$. Thus, starting from (359), we have

$$
\begin{aligned}
&1 - 2\mathrm{e}^{-sC_{1-s}(f(A)|E|P_{AE})} \\
&\ge 1 - 2c^{-s}\sum_e P_E(e)\sum_{a:P_{A|E}(a|e)\ge\frac{c}{M}} P_{A|E}(a|e)^{1-s}M^s \\
&\qquad - 2\cdot 2^{\frac{s}{1-s}}s^{\frac{s}{1-s}}(1-s)P_{AE}\Big\{(a,e) : P_{A|E}(a|e) < \frac{c}{M}\Big\} && (385)\\
&\ge 1 - 2c^{-s}P_{AE}\Big\{(a,e) : P_{A|E}(a|e) \ge \frac{c}{M}\Big\} \\
&\qquad - 2c\cdot 2^{\frac{s}{1-s}}s^{\frac{s}{1-s}}(1-s)P_{AE}\Big\{(a,e) : P_{A|E}(a|e) < \frac{c}{M}\Big\}. && (386)
\end{aligned}
$$

This completes the proof of (117). ■

Acknowledgements

The authors would like to acknowledge the Associate Editor (Prof. Aaron B. Wagner) and the anonymous reviewers for their extensive and useful comments during the revision process.

MH is partially supported by a MEXT Grant-in-Aid for Scientific Research (A) No. 23246071. MH is also partially supported by the National Institute of Information and Communication Technology (NICT), Japan. The Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

VYFT is partially supported an NUS Young Investigator Award (R-263-000-B37-133) and a Singapore Ministry of Education Tier 2 grant “Network Communication with Synchronization Errors: Fundamental Limits and Codes” (R-263-000-B61-112).

REFERENCES

[1] J. L. Carter and M. N. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, 18:143–154, 1979.

[2] Y. Liang, H. V. Poor, and S. Shamai. Information-theoretic security. *Foundations and Trends® in Communications and Information Theory*, 5(4–5):355–580, 2008.

[3] M. Bloch and J. Barros. *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[4] M. Iwamoto and J. Shikata. Information theoretic security for encryption based on conditional Rényi entropies. *Lecture Notes in Computer Science (Information Theoretic Security)*, 8317:103–121, 2014.

[5] S. Beigi and A. Gohari. Quantum achievability proof via collision relative entropy. *IEEE Trans. on Inform. Th.*, 60(12):7980–7986, Dec 2014.

[6] R. Impagliazzo, L. A. Levin, and M. Luby. Pseudo-random generation from one-way functions. In *In Proceedings of the 21st Annual ACM Symposium on Theory of Computing (STOC '89)*, pages 12–24. ACM Press, 1989.

[7] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer. Generalized privacy amplification. *IEEE Trans. on Inform. Th.*, 41(6):1915–1923, 1999.

[8] J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28:12–24, 1999.

[9] Y. Dodis and Y. Yu. Overcoming weak expectations. *Lecture Notes in Computer Science*, 7785:1–22, 2013.

[10] Y. Yao and Z. Li. Overcoming weak expectations via the Renyi entropy and the expanded computational entropy. *Lecture Notes in Computer Science*, 8317:162–178, 2014.

[11] J. Shikata. Design and analysis of information-theoretically secure authentication codes with non-uniformly random keys. *Cryptology ePrint Archive: Report 2015/250*, 2015.

[12] A. D. Wyner. The wire-tap channel. *The Bell System Technical Journal*, 54:1355–1387, 1975.

[13] R. G. Gallager. *Information Theory and Reliable Communication*. Wiley, New York, 1968.

[14] I. Csiszár and J. Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.

[15] V. Strassen. Asymptotische Abschätzungen in Shannons Informationstheorie. In *Trans. Third Prague Conf. Inf. Theory*, pages 689–723, Prague, 1962. http://www.math.cornell.edu/∼pmlut/strassen.pdf.

[16] M. Hayashi. Second-order asymptotics in fixed-length source coding and intrinsic randomness. *IEEE Trans. on Inform. Th.*, 54(10):4619–4637, 2008.

[17] M. Hayashi. Exponential decreasing rate of leaked information in universal random privacy amplification. *IEEE Trans. on Inform. Th.*, 57(6):3989–4001, 2011.

[18] M. Hayashi. Tight exponential analysis of universally composable privacy amplification and its applications. *IEEE Trans. on Inform. Th.*, 59(11):7728–7746, 2013.

[19] R. Ahlswede and I. Csiszár. Common randomness in information theory and cryptography–I: Secret sharing. *IEEE Trans. on Inform. Th.*, 39(4):1221–1132, 1993.

[20] T.-H. Chou, S. C. Draper, and A. Sayeed. Key generation using external source excitation: Capacity, reliability, and secrecy exponent. *IEEE Trans. on Inform. Th.*, 58(4):2455–2474, 2012.

[21] T.-H. Chou, V. Y. F. Tan, and S. C. Draper. The sender-excited secret key agreement model: Capacity, reliability and secrecy exponents. *IEEE Trans. on Inform. Th.*, 61(1):609–627, 2015.

[22] M. Hayashi. General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel. *IEEE Trans. on Inform. Th.*, 52(4):1562–1575, Apr 2006.

[23] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, Dipl. Phys. ETH, Switzerland, 2005.

[24] R. Renner and S. Wolf. Simple and tight bounds for information reconciliation and privacy amplication. In *Advances in Cryptology–ASIACRYPT 2005*, 2005.

[25] M. Hayashi. Large deviation analysis for quantum security via smoothing of Rényi entropy of order 2. *IEEE Trans. on Inform. Th.*, 60(10):6702–6732, 2014.

[26] M. Hayashi and T. Tsurumaru. More efficient privacy amplification with less random seeds via dual universal hash function. *IEEE Trans. on Inform. Th.*, 62(4):2213–2232, Apr 2016.

[27] J. Hou and G. Kramer. Effective secrecy: Reliability, confusion and stealth. *ArXiv 1311.1411*, 2013.

[28] J. Hou and G. Kramer. Informational divergence approximations to product distributions. In *13th Canadian Workshop on Inf. Theory (CWIT)*, pages 76–81, 2013.

[29] A. J. Pierrot and M. R. Bloch. Joint channel intrinsic randomness and channel resolvability. In *IEEE Information Theory Workshop (ITW)*, 2013.

[30] M. Bloch and J. N. Laneman. Strong secrecy from channel resolvability. *IEEE Trans. on Inform. Th.*, 59(12):8077–8098, 2013.

[31] T. S. Han, H. Endo, and M. Sasaki. Reliability and security functions of the wiretap channel under cost constraint. *IEEE Trans. on Inform. Th.*, 60(11):6819–6843, Nov 2014.

[32] M. B. Parizi and I. E. Telatar. On the secrecy exponent of the wire-tap channel. In *IEEE Information Theory Workshop (ITW)*, Jeju, South Korea, 2015.

[33] M. B. Parizi, E. Telatar, and N. Merhav. Exact random coding secrecy exponents for the wiretap channel. *Submitted to the IEEE Trans. on Inform. Th.*, 2016. `arXiv:1601.04276 [cs.IT]`.

[34] R. G. Gallager. Source coding with side information and universal coding. Technical report, MIT LIDS, 1976.

[35] I. Csiszár and P. Narayan. Secrecy capacities for multiple terminals. *IEEE Trans. on Inform. Th.*, 50(12):3047–3061, 2004.

[36] M. Hayashi. Security analysis of ε-almost dual universal$_2$ hash functions: smoothing of min entropy vs. smoothing of Rényi entropy of order 2. *IEEE Trans. on Inform. Th.*, 62(6):3451–3476, Jun 2016.

[37] T. Van Erven and P. Harremoës. Rényi divergence and Kullback-Leibler divergence. *IEEE Trans. on Inform. Th.*, 60(7):3797–3820, 2014.

[38] R. Sibson. Information radius. *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 14(2):149–160, 1969.
[39] S. Verdú. Alpha-mutual information. In *IEEE Information Theory and Applications Workshop*, San Diego, Feb 2015.
[40] M. Hayashi. Quantum wiretap channel with non-uniform random number and its exponent and equivocation rate of leaked information. *IEEE Trans. on Inform. Th.*, 61(10):5595–5622, 2015.
[41] M. Tomamichel and M. Hayashi. Operational interpretation of Renyi information measures via composite hypothesis testing against product and Markov distributions. *Submitted to the IEEE Trans. on Inform. Th.*, 2015. `arXiv:1511.04874 [cs.IT]`.
[42] S.-I. Amari and H. Nagaoka. *Methods of Information Geometry*. American Mathematical Society, 2000.
[43] N. Sharma and N. A. Warsi. Fundamental bound on the reliability of quantum information transmission. *Physical Review Letters*, 110(8):080501, 2013.
[44] S. Arimoto. Information measures and capacity of order α for discrete memoryless channels. In *Colloquia Mathematica Societatis János Bolya*, pages 41–52, Kestheley, Hungary, 1975.
[45] T. M. Cover. A proof of the data compression theorem of Slepian and Wolf for ergodic sources. *IEEE Trans. on Inform. Th.*, 21(3):226–228, 1975.
[46] U. Maurer and S. Wolf. Information-theoretic key agreement: From weak to strong secrecy for free. In *Proc. 19th Intl. Conf. Theory Appl. Cryptographic Tech.*, EUROCRYPT'00, pages 351–368. Springer-Verlag, 2000.
[47] T. S. Han. *Information-Spectrum Methods in Information Theory*. Springer Berlin Heidelberg, Feb 2003.
[48] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, 2nd edition, 1998.
[49] S. Verdú and I. Kontoyiannis. Optimal lossless data compression: Non-asymptotics and asymptotics. *IEEE Trans. on Inform. Th.*, 60(2):777–795, 2014.
[50] V. Kostina and S. Verdú. Fixed-length lossy compression in the finite blocklength regime. *IEEE Trans. on Inform. Th.*, 58(6):3309–3338, 2012.
[51] V. Y. F. Tan and O. Kosut. On the dispersions of three network information theory problems. *IEEE Trans. on Inform. Th.*, 60(2):881–903, 2014.
[52] M. Hayashi. Information spectrum approach to second-order coding rate in channel coding. *IEEE Trans. on Inform. Th.*, 55(11):4947–4966, 2009.
[53] V. Y. F. Tan. Asymptotic estimates in information theory with non-vanishing error probabilities. *Foundations and Trends® in Communications and Information Theory*, 11(1–2):1–184, 2014.
[54] Y. Altuğ and A. B. Wagner. Moderate deviations in channel coding. *IEEE Trans. on Inform. Th.*, 60(8):4417–4426, 2014.
[55] Y. Polyanskiy and S. Verdú. Channel dispersion and moderate deviations limits for memoryless channels. In *Proc. of Allerton Conference*, 2010.
[56] Y. Altuğ, A. B. Wagner, and I. Kontoyiannis. Lossless compression with moderate error probability. In *Proc. of Intl. Symp. on Inform. Th.*, Istanbul, Turkey, 2013.
[57] V. Y. F. Tan. Moderate-deviations of lossy source coding for discrete and Gaussian sources. In *Proc. of Intl. Symp. on Inform. Th.*, Cambridge, MA, 2012.
[58] V. Y. F. Tan, S. Watanabe, and M. Hayashi. Moderate deviations for joint source-channel coding of systems with Markovian memory. In *Proc. of Intl. Symp. on Inform. Th.*, Honolulu, HI, 2014.
[59] M. Tomamichel and V. Y. F. Tan. Second-order coding rates for channels with state. *IEEE Trans. on Inform. Th.*, 60(8):4427–4448, 2014.
[60] M. Sion. On general minimax theorems. *Pac. J. Math.*, 8(1):171–176, 1958.
[61] A. Ingber and Y. Kochman. The dispersion of lossy source coding. In *Proceedings of the Data Compression Conference (DCC)*, 2011. `arXiv:1102.2598 [cs.IT]`.
[62] M. Hayashi and V. Y. F. Tan. Asymmetric evaluations of erasure and undetected error probabilities. *IEEE Trans. on Inform. Th.*, 61(12):6560–6577, Dec 2015.
[63] S. Vembu and S. Verdú. Generating random bits from an arbitrary source: fundamental limits. *IEEE Trans. on Inform. Th.*, 41(5):1322–1332, 1995.
[64] W. Feller. *An Introduction to Probability Theory and Its Applications*. John Wiley and Sons, 2nd edition, 1971.
[65] T. Tao. *Topics in Random Matrix Theory*. American Mathematical Society, 2012.